BCG × SIA SEMICONDUCTOR INDUSTRY ASSOCIATION

STRENGTHENING THE GLOBAL SEMICONDUCTOR SUPPLY CHAIN IN AN UNCERTAIN ERA

APRIL 2021

ANTONIO VARAS / RAJ VARADARAJAN / JIMMY GOODRICH / FALAN YINUG

About Boston Consulting Group (BCG)

Boston Consulting Group (BCG) is a leading global management consulting firm, with offices in over 50 countries. BCG partners with leaders in business and society to tackle their most important challenges and capture their greatest opportunities. BCG was the pioneer in business strategy when it was founded in 1963. Today, we help clients with total transformation—inspiring complex change, enabling organizations to grow, building competitive advantage, and driving bottom-line impact.

About the Semiconductor Industry Association (SIA)

The Semiconductor Industry Association (SIA) is the voice of the semiconductor industry in the US, one of America's top export industries and a key driver of America's economic strength, national security, and global competitiveness. The semiconductor industry directly employs nearly a quarter of a million workers in the United States, and US semiconductor company sales totaled $208 billion in 2020. SIA represents 98 percent of the US semiconductor industry by revenue and nearly two-thirds of non-US chip firms. Through this coalition, SIA seeks to strengthen leadership of semiconductor manufacturing, design, and research by working with Congress, the Administration, and key industry stakeholders around the world to encourage policies that fuel innovation, propel business, and drive international competition.

About the Authors

Antonio Varas is a Senior Partner and Managing Director in the Silicon Valley office of Boston Consulting Group and is a core member of its Technology, Media & Telecommunications practice. You may contact him by email at varas.antonio@bcg.com.

Raj Varadarajan is a Senior Partner and Managing Director in the Dallas office of Boston Consulting Group and leads its Global Advantage practice in North America. You may contact him by email at varadarajan.raj@bcg.com.

Jimmy Goodrich is the Vice President, Global Policy at the Semiconductor Industry Association. You may contact him at jgoodrich@semiconductors.org.

Falan Yinug is the Director, Industry Statistics and Economic Policy at the Semiconductor Industry Association. You may contact him at fyinug@semiconductors.org.

Acknowledgments

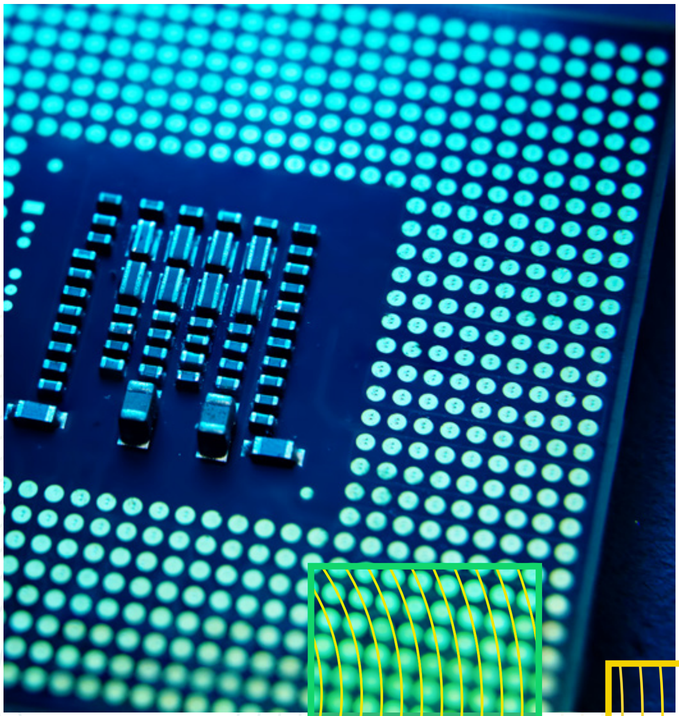
This report would not have been possible without the contributions of our BCG colleagues Ramiro Palma, Anoop Shah, Zhenyu Bo, Kaitlyn Alsup, Sohini Kar, Yeonsoo Lee, Ian Kleinman, and our SIA colleague Devi Keller.

Contents

Executive Summary

The widespread shortage of semiconductors that began in late 2020 highlighted how indispensable these specialized components are in today's economy. Semiconductors are used to power a vast array of electronic devices—everything from smartphones and cloud servers to modern cars, industrial automation, and critical infrastructure and defense systems.

The global structure of the semiconductor supply chain, developed over the past three decades, has enabled the industry to deliver continual leaps in cost savings and performance enhancements that ultimately made possible the explosion in information technology and digital services. In the past few years, however, several new factors have emerged that could put the successful continuation of this global model at risk. Addressing these vulnerabilities requires a combination of carefully designed actions from policymakers, including targeted incentives to encourage domestic production in order to address strategic gaps.



An Integrated Global Supply Chain

Semiconductors are highly complex products to design and manufacture. No other industry has the same high level of investment in both R&D (22% of annual semiconductor sales to electronic device makers) and capital expenditure (26%).

The need for deep technical know-how and scale has resulted in a highly specialized global supply chain, in which regions perform different roles according to their comparative advantages. (See Exhibit 1.) The US leads in the most R&D-intensive activities—electronic design automation (EDA), core intellectual property (IP), chip design, and advanced manufacturing equipment—owing to its world-class universities, vast pool of engineering talent and market-driven innovation ecosystem. East Asia is at the forefront in wafer fabrication, which requires massive capital investments supported by government incentives as well as access to robust infrastructure and skilled workforce. China is a leader in assembly, packaging and testing, which is relatively less skill- and capital-intensive, and is investing aggressively to expand throughout the value chain.

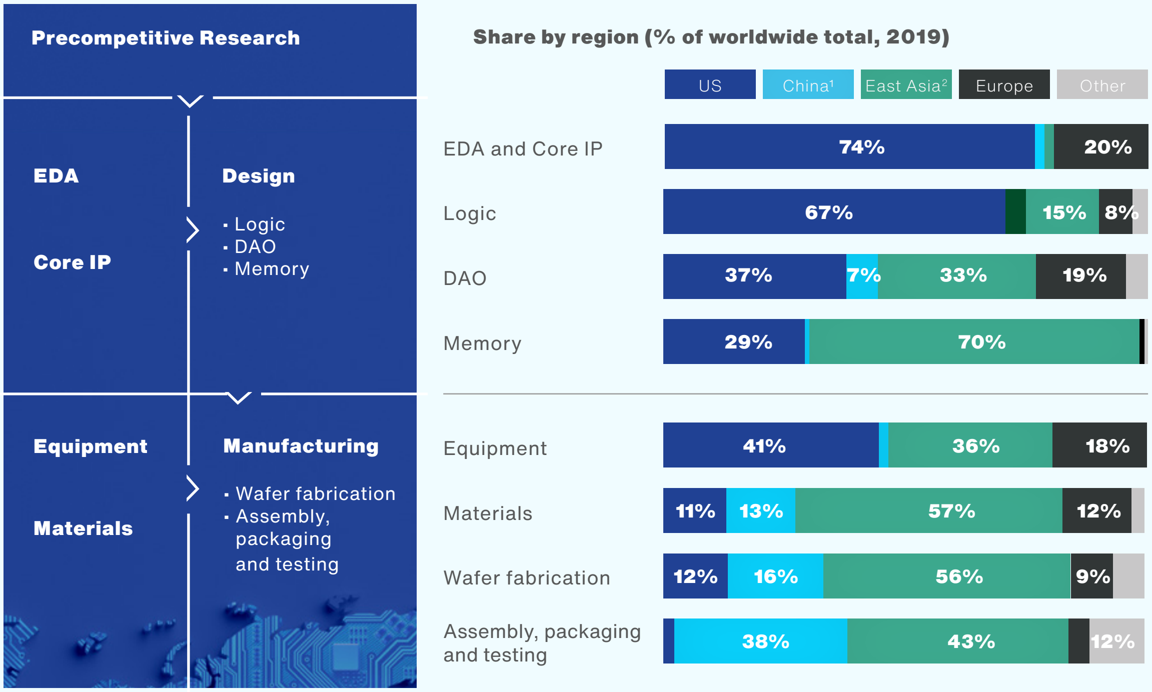
All countries are interdependent in this integrated global supply chain, relying on free trade to move materials, equipment, IP, and products around the world to the optimal location for performing each activity. In fact, semiconductors are the world's fourth-most-traded product after only crude oil, refined oil, and cars.

This global structure delivers enormous value. A hypothetical alternative with parallel, fully "self-sufficient" local supply chains in each region to meet its current levels of semiconductor consumption would have required at least $1 trillion in incremental upfront investment, resulting in a 35% to 65% overall increase in semiconductor prices and ultimately higher costs of electronic devices for end users.

EXHIBIT 1

The global semiconductor supply chain based on geographic specialization has delivered enormous value for the industry

Semiconductor Supply Chain



Cost savings vs. fully localized "self-sufficient" supply chains:

$0.9-1.2T
avoided upfront investment

$45-125B
annual cost efficiencies

35-65%
enabled reduction in semiconductor prices

Source: BCG analysis
Note: DAO = discrete, analog, and other (including optoelectronics and sensors); EDA = electronic design automation; OSAT = outsourced assembly and test
1. Mainland China 2. East Asia includes South Korea, Japan, and Taiwan

Risks and Vulnerabilities

Over the next ten years, the industry will need to invest about $3 trillion in R&D and capital expenditure globally across the value chain in order to meet the increasing demand for semiconductors. Industry participants and governments must collaborate to continue facilitating worldwide access to markets, technologies, capital, and talent, and make the supply chain more resilient.

While geographic specialization has served the industry well, it also creates vulnerabilities that each region needs to assess in a manner specific to its own economic and security considerations. There are more than 50 points across the supply chain where one region holds more than 65% of the global market share, although the level of risk associated with each of these varies. Manufacturing emerges as a major focal point when it comes to the resilience of the global semiconductor supply chain. About 75% of semiconductor manufacturing capacity, as well as many suppliers of key materials—such as silicon wafers, photoresist, and other specialty chemicals—are concentrated in China and East Asia, a region significantly exposed to high seismic activity and geopolitical tensions. Furthermore, all of the world's most advanced semiconductor manufacturing capacity—in nodes below 10 nanometers—is currently located in South Korea (8%) and Taiwan (92%). These are single points of failure that could be disrupted by natural disasters, infrastructure shutdowns, or international conflicts, and may cause severe interruptions in the supply of chips.

Besides the risks associated with concentration in certain geographic locations, geopolitical tensions may result in export controls that impair access to critical providers of essential technology, tools, and products that are clustered in certain countries. Such controls could also restrict access to important end markets, potentially resulting in a significant loss of scale and compromising the industry's ability to sustain the current levels of R&D and capital intensity.

The solution to these challenges is not the pursuit of complete self-sufficiency through large-scale national industrial policies with a staggering cost and questionable execution feasibility. Instead, the semiconductor industry needs nuanced targeted policies that strengthen supply chain resilience and expand open trade, while balancing the needs of national security.

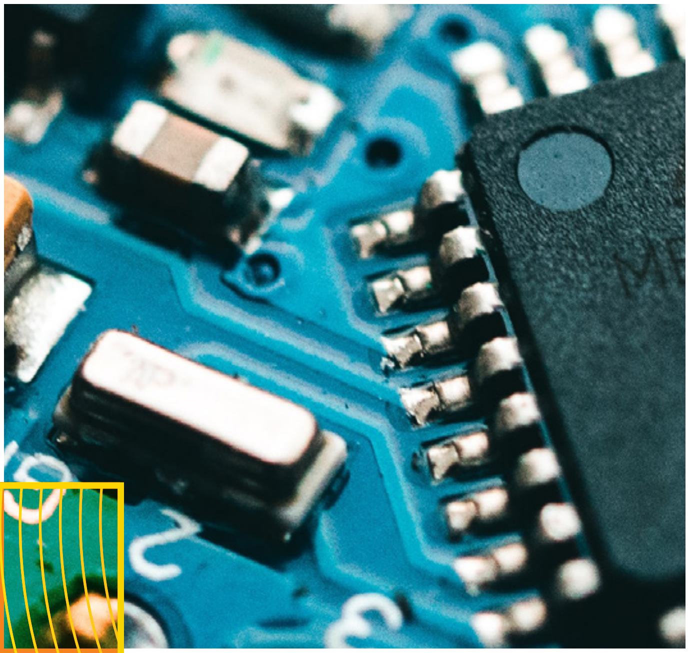
To address the risk of major global supply disruptions, governments should enact market-driven incentive programs to achieve a more diversified geographical footprint, which should include building additional manufacturing capacity in the US, as well as expanding the production sites and sources of supply for some critical materials. In our **previous report** "Government Incentives and US Competitiveness in Semiconductor Manufacturing," we found that a $50 billion incentive program would establish the US as an attractive location for semiconductor manufacturing. Our analysis shows that such a program could enable the construction of 19 advanced fabs for logic, memory, and analog semiconductors over the next ten years, doubling the number expected if no action is taken.

This new capacity would be instrumental to address major vulnerabilities in the supply chain. For example, it would allow the US to maintain a minimum viable manufacturing capacity in the leading nodes to meet domestic demand for the advanced logic chips used in national security systems, aerospace, and critical infrastructure. In contrast, we estimate that a goal of complete manufacturing self-sufficiency—seeking to cover the total US semiconductor consumption with onshore capacity—would require over $400 billion in government incentives and cost more than one trillion dollars over ten years.

In setting policies to promote supply chain resilience, governments must guarantee a level global playing field for domestic and foreign firms alike, as well as strong protection of IP rights. They must also take steps to further promote global trade and international collaboration on R&D and technology standards. In parallel, policymakers need to step up efforts to stimulate basic research and address the shortage of talent threatening to constrain the industry's ability to maintain its innovation pace. To that end, further public investment in science and engineering education is needed, as well as immigration policies that enable leading global semiconductor clusters to attract world-class talent.

In addition, governments with significant national security concerns should establish a clear and stable framework for targeted controls on semiconductor trade that avoid broad unilateral restrictions on technologies and vendors.

Such well-modulated policy interventions would preserve the benefits of scale and specialization in today's global supply chain structure. This would ensure that the industry can extend its ability to deliver the continual improvements in semiconductor performance and cost that will make the promise of transformative technologies such as AI, 5G, IoT, and autonomous electric vehicles a reality in this decade.



Introduction

Today's mobile phone users likely do not put much thought into the complex cross-border collaboration in research, development, design, and manufacturing amongst hundreds of firms that enable them to access their favorite content over a high-speed wireless network. Consumers, however, benefit from the global coordination throughout the deep and complex electronics industry in the form of accelerated innovation cycles that deliver new technology features at lower prices. The backbone of this globally integrated digital economy is the semiconductor supply chain.

Over the past three decades the semiconductor industry has experienced rapid growth and delivered enormous economic impact. The semiconductor market grew at a 7.5% compound annual growth rate from 1990 to 2020, outpacing the 5% growth of global GDP during that time. The performance and cost improvements delivered by the semiconductor industry made possible the evolution from mainframes to PCs in the 1990s, the client-server architecture underpinning the Web and online services in the 2000s, and then the advent of the smartphone as the computer in everyone's pocket in the 2010s.

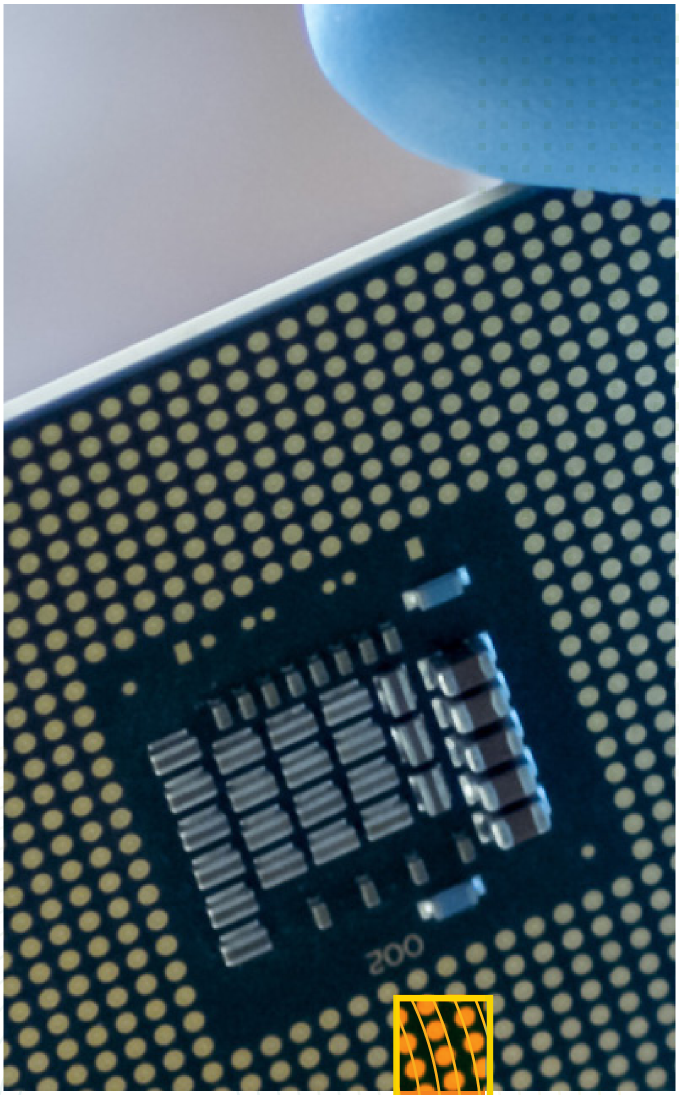
These innovations have created tremendous economic growth: an estimated additional $3 trillion in global GDP from 1995 to 2015 has been directly linked with semiconductor innovation, with an incremental $11 trillion in indirect impact[1]. Going forward, further advancements in semiconductor technology will be essential to enable a new wave of transformative technologies, including Artificial Intelligence (AI), 5G, autonomous electric vehicles or Internet of Things (IoT) solutions deployed at scale with a myriad of smart connected devices.

This economic impact has been made possible by relentless accelerated improvement in semiconductor technology. Since the invention of the integrated circuit back in 1958, the number of transistors per wafer for a logic chip has increased by a factor of about 10 million, yielding a 100,000-fold gain in processor speed and a cost reduction of more than 45% per year for comparable performance. Coupled with engineering innovations such as advanced packaging and materials technology, this has allowed electronic device

---

1. IHS report "*Moore's Law Impact*", May 2015.

makers to create devices with exponentially more computational power in increasingly smaller form factors. As an illustration, today's smartphones have more computer power than the mainframe computers that NASA used to send Apollo 11 to the moon in 1969. Today's smartphones also contain more memory for storage than a data center server in 2010. Similarly, advances in analog semiconductors have also made possible dramatic improvements in quality and speed of wireless communications under successive generations of cellular technology, leading to the recently introduced 5G.



The semiconductor supply chain is the backbone of the digital economy

Semiconductors are highly complex products fabricated from highly advanced manufacturing processes. Improvements often require breakthroughs in hard science that take many years to achieve. The fast pace of innovation in the semiconductor industry is the result of enormous investments and a sophisticated global value chain and research infrastructure integrated by highly specialized companies and institutions distributed across the world.

Specialization across the supply chain allows the deep focus required to innovate, often pushing the boundaries of science. There are more than 30 types of semiconductor product categories, each optimized for a particular function in an electronic subsystem. Developing a modern chip requires deep technical expertise in both hardware and software, and relies on advanced design tools and intellectual property (IP) provided by specialized firms. Fabrication then typically requires as many as 300 different inputs, including raw wafers, commodity chemicals, specialty chemicals, and bulk gases. These inputs are processed by more than 50 classes of highly engineered precision equipment. Most of this equipment, such as lithography and metrology tools, incorporates hundreds of technology subsystems such as modules, lasers, mechatronics, control chips, and optics. The highly specialized suppliers involved in semiconductor design and fabrication are often based in different countries. Chips then zigzag across the world in a global journey.

This report seeks to provide an understanding of the complex global semiconductor supply chain, how it supports the industry's continuous technology innovation, and how it ultimately benefits consumers and enables our economy through better technology at lower prices. We also identify a number of risks that could affect the industry's ability to continue delivering exponential performance and cost improvements and discuss ways to address them.

Overview of the Semiconductor Supply Chain

Understanding semiconductors: what they are and what they are used for

Semiconductors are highly specialized components that provide the essential functionality for electronic devices to process, store and transmit data. Most of today's semiconductors are integrated circuits, also referred to as "chips". A chip is a set of miniaturized electronic circuits composed of active discrete devices (transistors, diodes), passive devices (capacitors, resistors) and the interconnections between them, layered on a thin wafer of semiconductor material, typically silicon. Modern chips are tiny, packing billions of electronic components in an area as small as only a few square millimeters.

While industry taxonomies typically describe more than 30 types of product categories, semiconductors can be classified into three broad categories:

1 Logic (42% of industry revenues)

These are integrated circuits functioning on binary codes (0 and 1) that serve as the fundamental building blocks or "brains" of computing:

Microprocessors are logic products such as central processing units (CPUs), graphics processing units (GPUs) and application processors (APs) that process fixed instructions stored on memory devices to execute complex computing operations. Applications include processors for mobile phones, personal computers, servers, AI systems, and supercomputers.

General purpose logic products such as Field Programmable Gate Arrays (FPGAs) do not contain any pre-fixed instructions, allowing a user to program custom logic operations.

Microcontrollers (MCUs) are small computers on a single chip. A microcontroller contains one or more processor cores along with memory and programmable input/output peripherals. MCUs perform basic computing tasks in myriads of electronic products such as cars, industrial automation equipment or consumer appliances.

Connectivity products, such as cellular modems, WiFi or Bluetooth chips or Ethernet controllers, allow electronic devices to connect to a wireless or wired network to transmit or receive data.

2 Memory (26% of industry revenues)

These are semiconductors used for storing information necessary to perform any computation. Computers process information stored in their memory, which consists of various data storage or memory devices. Two most commonly used semiconductor memories in use today are Dynamic Random-Access Memory (DRAM) and NAND memory:

DRAM is used to store the data or program code needed by a computer processor to function. It is typically found in personal computers (PCs) and servers. Smartphones are also growing the DRAM content they require, and there is also increasing need for DRAM in automobile electronics applications such as advanced driver-assistance systems (ADAS).

NAND is the most common type of flash memory. Unlike DRAM, it does not need power to retain data, so it is used for permanent storage. Typical applications include solid state drives (SSDs) used as laptop hard drives or secure digital (SD) cards used in portable devices.

3 Discrete, Analog, and Other (DAO) (32% of the industry revenues)

These are semiconductors that transmit, receive, and transform information dealing with continuous parameters such as temperature and voltage:

Discrete products include diodes and transistors that are designed to perform a single electrical function.

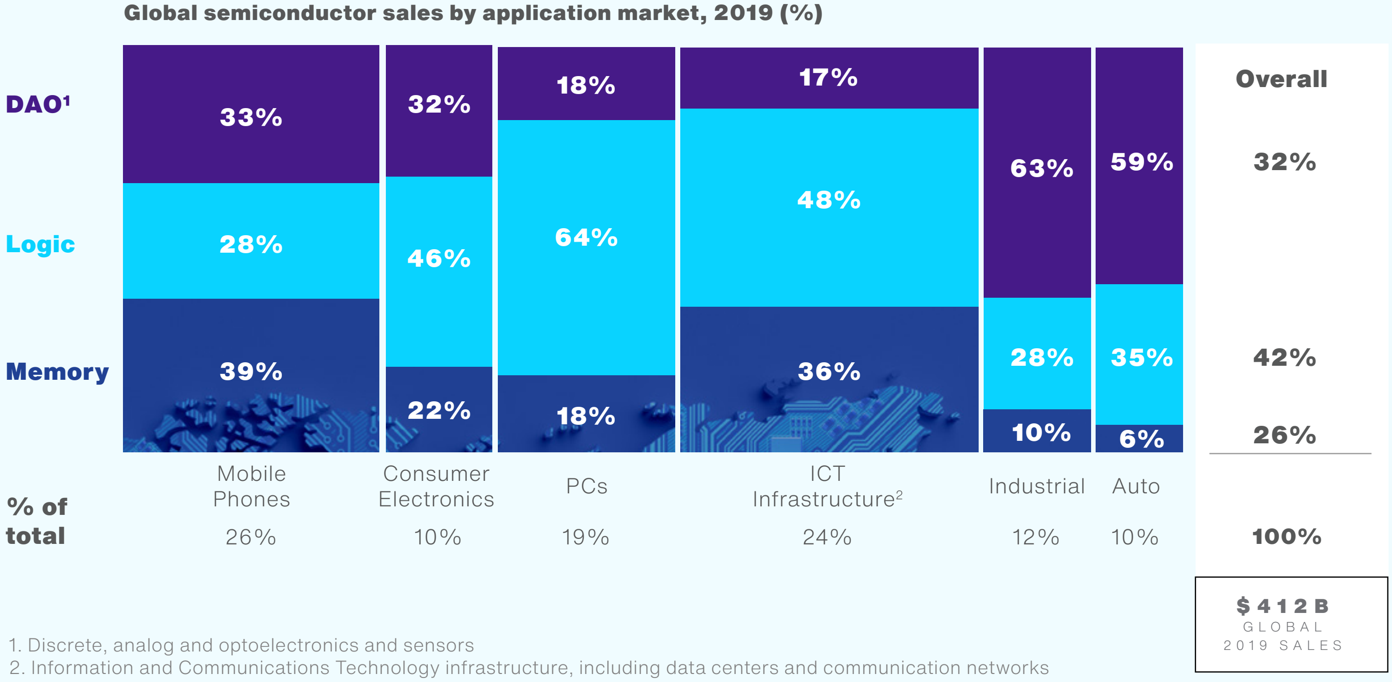
Analog products include voltage regulators and data converters that translate analog signals from sources such as voice into digital signals. This category also includes power management integrated circuits found in any type of electronic device, and radio frequency (RF) semiconductors that enable smartphones to receive and process the radio signals coming from the base stations of cellular networks.

Other products include optoelectronics, such as optical sensors to sense light used in cameras, as well as a wide variety of non-optical sensors and actuators that can be found in all sorts of Internet of Things devices.

Semiconductors are used in all types of electronic devices across multiple applications spanning the major sectors of the economy (Exhibit 2). Each of these application markets requires semiconductors from all three broad categories described above. For example, mobile phones have practically as much DAO content (essential for features such a cellular connectivity, camera and power consumption management) as logic content (which includes the microprocessors that provide increasing computing power with every new phone generation) and memory (for storage of digital content on the device). Approximately 65% of global semiconductor revenues are from general-purpose components that are used across multiple applications.

EXHIBIT 2

All types of semiconductors are indispensable in today's economy, powering all sorts of electronic devices



1. Discrete, analog and optoelectronics and sensors
2. Information and Communications Technology infrastructure, including data centers and communication networks
Sources: SIA WSTS, Gartner

Demand for semiconductors is highly global. The share of the global semiconductor demand that comes from each region is different depending on how the point of origin of demand is defined. While semiconductors are typically sourced by the electronic device makers to build their products, ultimately semiconductor demand is driven by the end users who purchase those devices. This is why, from a geographic standpoint, there are three different ways of measuring the origin of semiconductor demand, with reference to alternative points in the global electronics supply chain:

A

Location of the headquarters of the electronic device makers. These firms are the customers of the chip companies, purchasing the semiconductors that go into their devices. Electronic device makers – commonly referred to by the term original equipment manufacturers (OEMs) – typically design their products and decide which components to use from which vendor. For example, under this approach the semiconductors that go into a smartphone developed by a company headquartered in the US would be computed as US demand, even though the product may be physically built in another country.

B

Location where the device is manufactured/assembled: OEMs often do not manufacture their devices in the same country where their headquarters are located or where the engineering team that designed the device is based. Instead, the devices are typically assembled in a manufacturing plant located in a different or many different countries, often by other firms commonly referred to as original device manufacturers (ODMs) or electronic manufacturing services (EMS). This is the location where the finished semiconductors need to be physically shipped to. For example, with this approach the chips going into a smartphone designed by a US company but actually manufactured by a Taiwanese contractor in a plant located in mainland China, would be computed as Chinese demand.

C

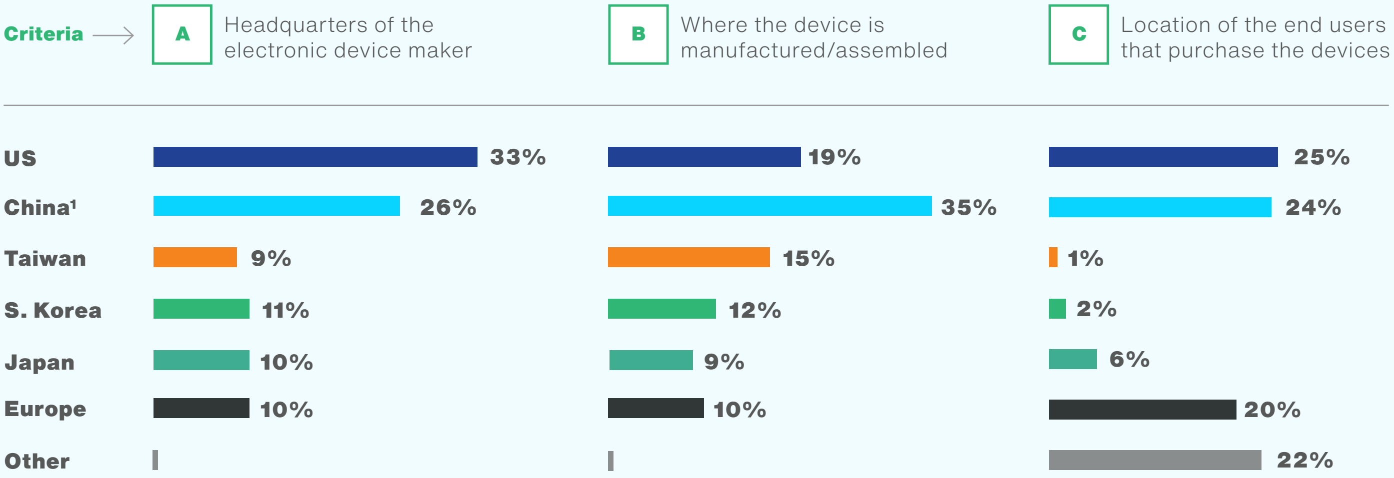
Location of the end users that purchase the electronic devices. Given that semiconductors are components, semiconductor demand is ultimately driven by sales of electronic devices to end users, both consumers and businesses. In our example, the value of the chips contained in smartphones designed by a US company but assembled in China would be distributed across all the countries in the world where these smartphones are sold to consumers.

Exhibit 3 shows the geographic breakdown of the global semiconductor demand using these three alternative lenses: the shares of countries or regions are quite different depending on the criteria. But none of the three potential approaches is deemed to be the “correct” answer – they just reflect the diverse roles that countries/regions play in the broader electronics industry.

EXHIBIT 3

Alternative views of geographic origin of semiconductor demand

Global semiconductor sales by geographic area, 2019 (%)



1. Mainland China
Sources: BCG analysis with data from SIA WSTS, Gartner, IDC

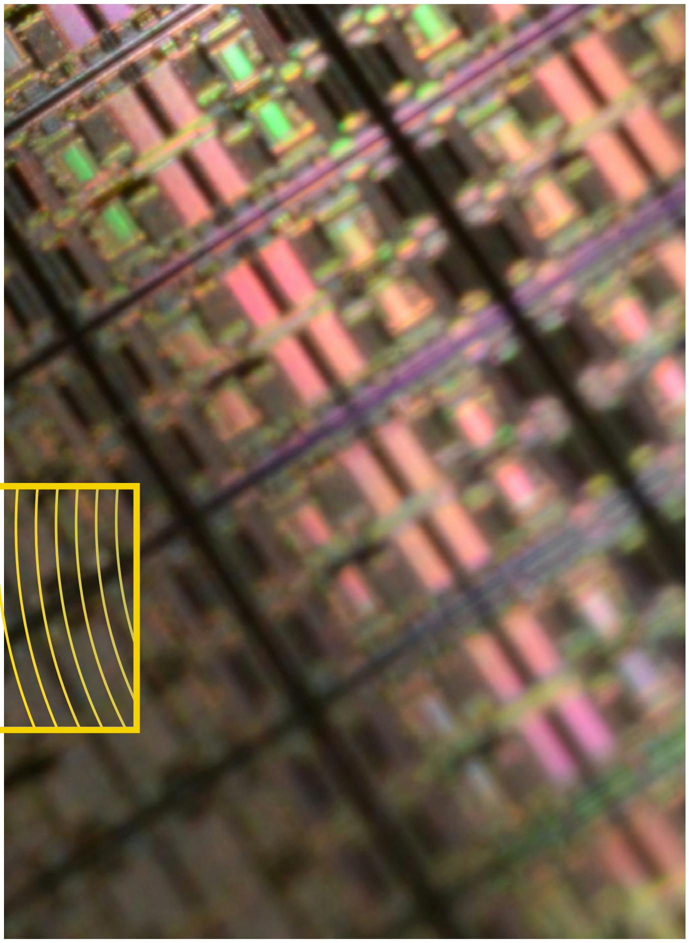
Considering the location of the electronic device makers (criterion A in the exhibit), the US still drives 33% of the total global semiconductor demand, the highest participation across all regions.

The US and China are the largest semiconductor markets, each accounting for 25% of global consumption

In 2019 US-based firms collectively accounted for about 45% market share in the large PCs and information and communications infrastructure (which includes data centers and network equipment) application markets, and a 30% market share in smartphones and industrial equipment. China has emerged as the clear second region, after tripling its participation in the past 10 years. The rise of China as a large source of semiconductor demand has been driven by the strength of its local firms in smartphones, PCs and consumer electronics: companies such as Huawei, Lenovo, Xiaomi and Oppo/Vivo not only sell their products for the domestic Chinese market, but also are major competitors in other markets.

China is the top region under the end electronics device manufacturing/assembly location criteria (criterion B in the exhibit), reflecting its strength in electronics manufacturing, particularly smartphones and consumer electronics products. As the world's main manufacturing hub, China was the destination for approximately 35% of total global chips sales in 2019. But many of the chips that enter China through this intermediate step are not ultimately consumed as products bought by Chinese end users, and instead are re-shipped overseas as components of made-in-China devices exported to other countries.

Focusing on where the devices are effectively sold to end users (criterion C in the exhibit) shows where semiconductor demand is ultimately coming from. Based on available market data for the different types of applications, we estimate that the value of the semiconductor content included in the devices bought by Chinese consumers and businesses represented approximately 24% of the global semiconductor revenues in 2019, practically at par with the US (25%) and a few percentage points above Europe (20%). However, the China's share of global semiconductor consumption is expected to continue to increase in the next 5 years, as analysts forecast that the growth of the Chinese domestic market will outperform the rest of the world by an average of 4-5 percentage points in most electronic device categories.

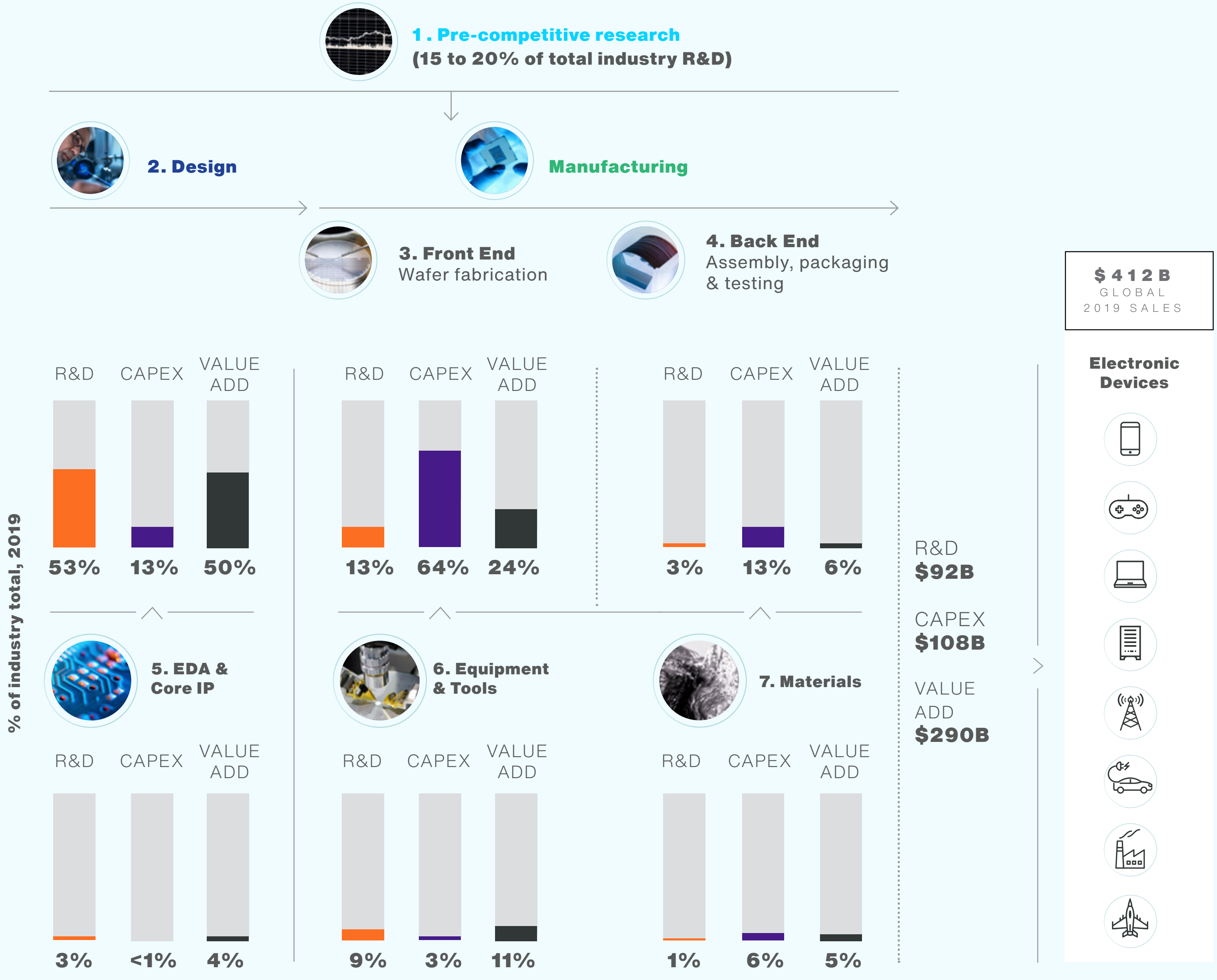


The structure of the semiconductor value chain

The industry value chain involved in the creation and production of any semiconductor is extraordinarily complex and globalized. At a high level, it consists of four broad steps, supported by a specialized ecosystem of materials, equipment and software design tools and core IP suppliers (Exhibit 4):

EXHIBIT 4

The semiconductor value chain includes seven differentiated activities



Sources: BCG analysis using data from Capital IQ (company financial reports) and Gartner (total market sizes)

Pre-competitive research
15-20% of overall industry R&D

Pre-competitive research is aimed at identifying the fundamental materials and chemical processes to seed the innovations in design architectures and manufacturing technology that will enable the next commercial leaps in computing power and efficiency. It is typically basic research in science and engineering, the results of which ordinarily are published and shared broadly within the scientific community, as distinguished from proprietary research and from industrial development, design and production. Pre-competitive basic research is qualitatively different from industry R&D: both types are complementary, not redundant[2]. In fact, pre-competitive research has been found to stimulate and attract industry R&D[3].

The average length of time between when a new technological approach is introduced in a research paper and when it hits widescale commercial manufacturing is estimated to be about 10-15 years, but it could be much longer than that for scientific breakthroughs that enable the current leading edge technologies. For example, Extreme Ultra-Violet (EUV) technology that is fundamental for the most advanced semiconductor manufacturing nodes took almost four decades from the early concept demos to its commercial implementation in fabs.

While there is no available data for the semiconductor industry, basic research typically accounts for 15-20% of the overall R&D investment in most leading countries. For example, in the US it has remained stable over time at 16-19% of the total R&D. Basic research is performed by a global network of scientists from private corporations, universities, government-sponsored national labs and other independent research institutions that collaborate in joint research efforts.

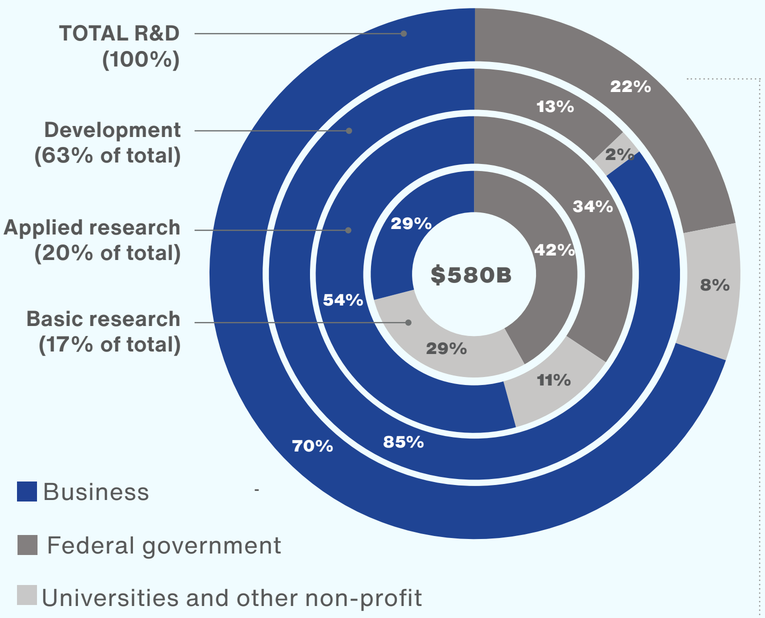
In particular, governments have a very significant role in advancing basic research. A prior study of federal R&D by the Semiconductor Industry Association (SIA) identified 8 major semiconductor technology breakthroughs that emerged out of government-sponsored research programs[4]. For example, the gallium arsenide (GaAs) transistors that enable smartphones to establish a wireless communication link to cellular towers was created in the Microwave and Millimeter Wave Integrated Circuit (MIMIC) program of the Department of Defense in the late 1980s.

The analysis of the total R&D investment in the US across all sectors provides some insights on the magnitude and distinct profile of pre-competitive research. Based on data compiled by the National Science Foundation, the US federal government is the main contributor to basic research with 42% of the investment in 2018. An additional 30% was funded by state governments, universities and other non-profit research institutions, and the remaining 28% came from companies. In contrast, the share of private companies in applied research and development – which typically follow after breakthroughs in basic research – was close to 80%

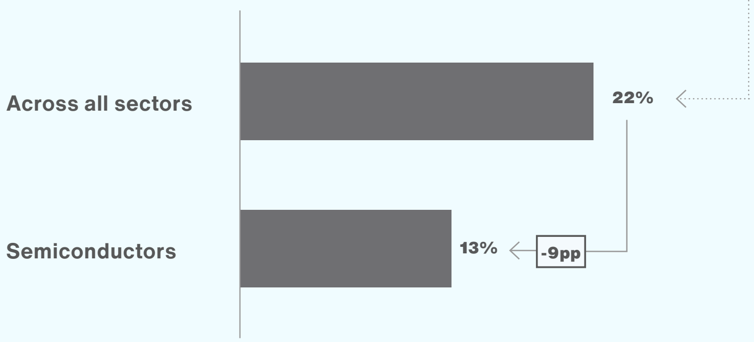
EXHIBIT 5

US illustrates the critical role of governments in basic research, although contribution to semiconductor sector seems trailing behind

Total US R&D investment across all sectors, 2018



Comparison of US federal government share in total R&D investment, 2018



Sources: BCG analysis with US National Science Foundation and OECD data, SIA

---

2. For a good illustration of the differentiated role of pre-competitive research and industry R&D, see the report by the United States Government Accountability Office on Nanomanufacturing, January 2014. 3. For more information on the positive relationship between federal and industry R&D and how increased federal R&D funding "crowds in" industry R&D spending, see the supplemental materials of the report by the Semiconductor Industry Association, Sparking Innovation, June 2020. 4. See Appendix C in Semiconductor Industry Association, Sparking Innovation, June 2020.

Funding for semiconductor basic research in the US seems to be trailing well behind the growth in applied research and development. The SIA study mentioned above found that the overall US federal government investment in semiconductor-related R&D – including basic research, applied research and development – accounted for just 13% of the total US semiconductor R&D in 2018. This percentage is significantly below the 22% share of federal government funding in the total US R&D spend across all sectors. In fact, while US private investment in semiconductor R&D as a percentage of GDP has increased nearly 10-fold over the last 40 years, federal investment has remained flat. Given the leading role that the US currently has in the most R&D-intensive activities across the semiconductor value chain, the impact of this gap in the funding of basic research may go beyond the relative competitiveness of US firms and create a risk for the overall industry's ability to maintain its historical pace of innovation.

In contrast, China is committing large sums to pre-competitive research as part of its effort to build a strong domestic semiconductor industry. In the last 20 years China has been closing the gap with the US on overall R&D spending. According to data from the Organization for Economic Cooperation and Development (OECD), in 2018 China was the world's second biggest spender on R&D in absolute terms: its total R&D investment was just 5% below the US in Purchasing Power Parity terms. However, just about 5-6% of Chinese R&D spend currently goes to basic research, significantly below all other countries with high investment levels in R&D.

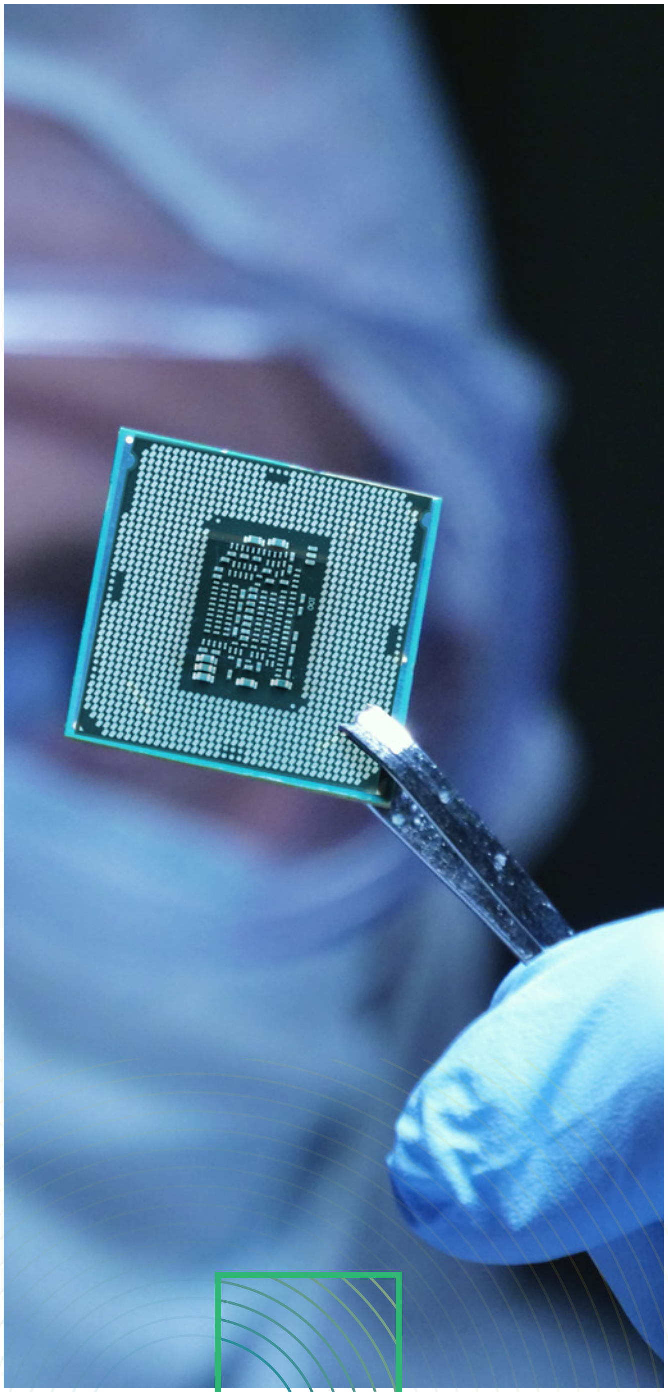
China's new five-year plan for 2021-25 announced in March clearly establishes boosting basic research as a critical priority. Central government spending on basic research will increase 11% in 2021, well above the 7% planned for the overall R&D investment and the 6% target for GDP growth. Semiconductors has been designated as one of the seven areas that will be given priority in terms of funding and resources.

Chip design
53% of industry R&D
13% of industry capex
50% of the value added

Firms involved in design develop the nanometer-scale integrated circuits which perform the critical tasks that make electronic devices work, such as computing, storage, connectivity to networks, and power management. Design relies on highly advanced electronic design automation (EDA) software and reusable architectural building blocks ("IP cores"), and in some cases also outsourced chip design services provided by specialized technology suppliers.

Semiconductors are highly complex products to design and manufacture

Design activity is largely knowledge- and skill-intensive: it accounts for 65% of the total industry R&D and 53% of the value added. Indeed, firms focusing on semiconductor design typically invest 12 to 20% of their annual revenues in R&D. Development of modern complex chips, such as the "system-on-chip" (SoC) processors that power today's smartphones, requires several years of effort by a large team of hundreds of engineers, sometimes leveraging external IP and design support services. Development costs have been rising rapidly as chips have become increasingly complex. The total development cost of a new state-of-the-art system-on-chip for a flagship smartphone, including the specialized blocks required to process audio, video or provide high-speed wireless connectivity, could well exceed $1 billion. Derivatives that reuse a significant portion of a prior design or new simpler chips that can be manufactured in mature nodes would cost just $20 million to $200 million to develop.



Wafer fabrication

(front end manufacturing)
13% of industry R&D
64% of industry capex
24% of the value added

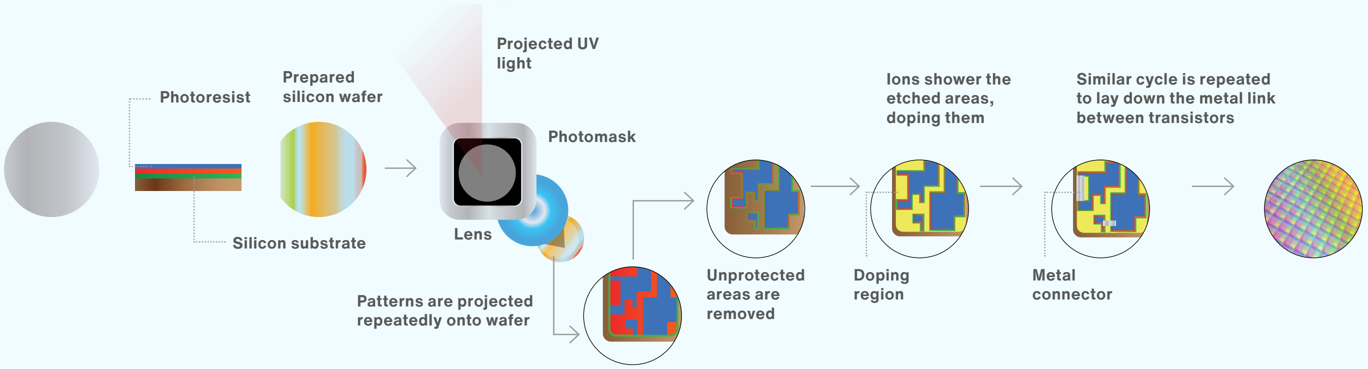
Highly specialized semiconductor manufacturing facilities, typically called “fabs”, print the nanometer-scale integrated circuits from the chip design into silicon wafers. Each wafer contains multiple chips of the same design. The actual number of chips per wafer depends on the size of the specific chip: it could vary between a hundred of the large, complex processors that power computers or smartphones, to hundreds of thousands for small chips intended to perform a simple function.

The fabrication process is intricate and requires highly specialized inputs and equipment to achieve the needed precision at miniature scale. Integrated circuits are built in cleanrooms, designed to maintain sterile conditions to prevent contamination by particles in the air that could alter the properties of the materials that form the electronic circuits. For comparison, the ambient outdoor air in a typical urban area contains 35,000,000 particles of 0.5 micron or bigger in size for each cubic meter, while a semiconductor manufacturing cleanroom permits absolutely zero particles of that size.

Depending on the specific product, there are 400 to 1,400 steps in the overall manufacturing process semiconductor wafers. The average time to fabricate finished semiconductor wafers, known as the cycle time, is about 12 weeks, but it can take up to 14-20 weeks to complete for advanced processes. It utilizes hundreds of different inputs, including raw wafers, commodity chemicals, specialty chemicals as well as many different types of processing and testing equipment and tools, across a number of stages (Exhibit 6). These steps are often repeated many hundreds of times, depending on the complexity of the desired set of electronic circuits.

EXHIBIT 6

Overview of the wafer fabrication process



Silicon wafer	1 **Oxidation and coating**	2 **Lithography**	3 **Etching**	4 **Doping**	5 **Metal deposition & etching**	**Completed wafer**
The silicon wafers start out blank and pure in a non-conductive state	Layers of insulating and conducting materials are applied on a silicon wafer. The wafer is then covered by a uniform coat of photoresist	The integrated circuit patterns specified in the design are mapped onto a glass plate called a photomask. An ultraviolet (UV) light is applied to transfer the patterns to the photoresist material coating on the wafer surface. The photoresist that was exposed to the light can now be chemically removed	Areas of the silicon wafer unprotected by photoresist are removed and cleaned by gases or chemicals	The wafer is showered with ionic gases that modify the properties of the new layer by adding a known quantity of impurities, such as boron and arsenic. Subsequent annealing will diffuse these impurities to a more uniform density	A similar process is used to lay down the metal links between transistors	Each completed wafer contains hundred of identical integrated circuits. The wafers are sent to the back-end manufacturing processes (assembly, packaging & testing)

Steps 1-4 are repeated hundreds of times with different chemicals to create more layers, depending on the desired circuit features

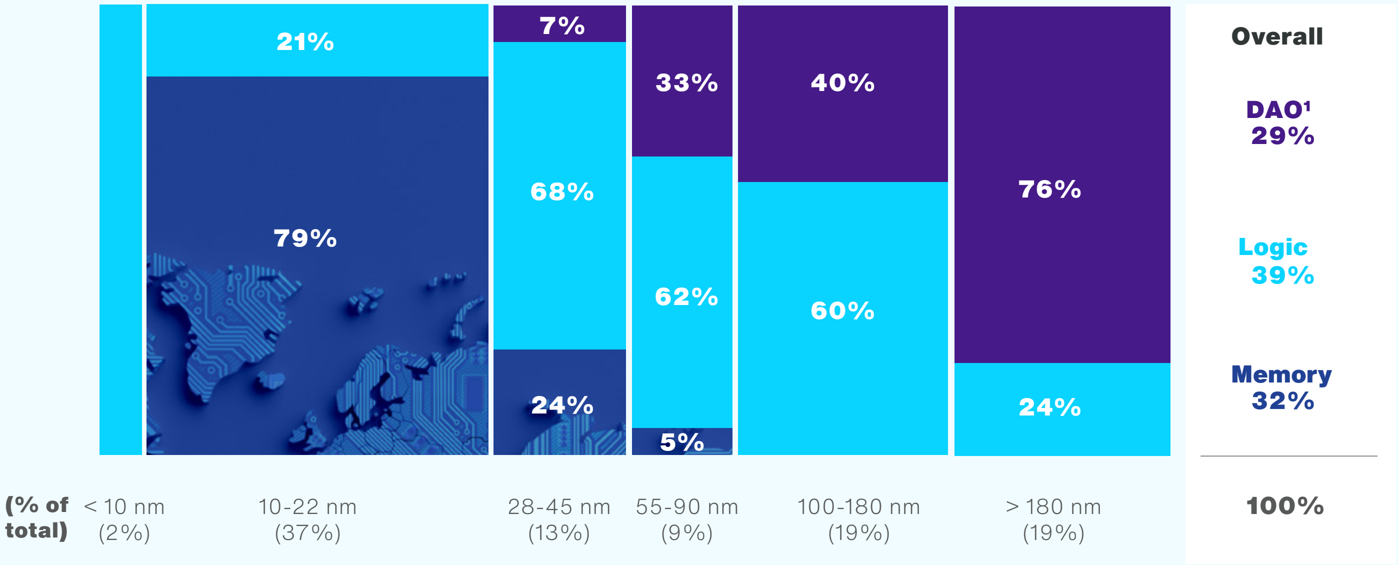
Advances in manufacturing process technology are typically described by referring to “nodes”. The term “node” is meant to refer to the size in nanometers of the transistor gates in the electronic circuits, although over time it has lost its original meaning and has become an umbrella term to designate both smaller features and also different circuit architectures and manufacturing technologies. Generally, the smaller the node size, the more powerful the chip, as more transistors can be placed on an area of the same size. This is the principle behind “Moore’s Law”, a key observation and projection in the semiconductor industry that states that the number of transistors on a logic chip doubles every 18 to 24 months. Moore’s Law has underpinned the relentless pace of simultaneous improvement in performance and cost for processors since 1965. Today’s advanced processors found in smartphones, computers, gaming consoles and data center servers are manufactured on 5 to 10-nanometer nodes. Commercial chip manufacturing using 3-nanometer process technology is expected to begin around 2023.

While logic and memory chips used for digital applications greatly benefit from the scaling in transistor size associated with smaller nodes, other types of semiconductors – particularly those in the DAO group described above—do not achieve the same degree of performance and cost benefits by migrating to ever smaller nodes, or simply use different types of circuits or architectures that would not work at more miniaturized scales. As a result, today wafer manufacturing still takes place across a wide range of nodes from the current “leading node” at 5 nanometers used for advanced logic to the legacy nodes above 180 nanometers used for discrete, optoelectronics, sensors, and analog semiconductors. In fact, only 2% of the global capacity is currently on nodes below 10 nanometers (Exhibit 7).

EXHIBIT 7

Semiconductor manufacturing utilizes capacity across a wide range of nodes

Global manufacturing capacity by node and semiconductor type, 2019
(% of 8” equivalent wafers per month)



1. Discrete, analog and optoelectronics and sensors
Sources: BCG analysis based on SEMI data

Front-end manufacturing is highly capital intensive due to the scale and complex equipment needed to produce semiconductors. A state-of-the-art semiconductor fab of standard capacity requires roughly $5 billion (for advanced analog fabs) to $20 billion (for advanced logic and memory fabs) of capital expenditure, including land, building, and equipment. This is significantly higher than, for example, the estimated cost of a next-generation aircraft carrier ($13 billion) or a new nuclear power plant ($4 billion to $8 billion)[5]. Capital expenditure of firms focusing on semiconductor manufacturing typically amounts to 30 to 40% of their annual revenues. As a result, wafer fabrication accounts for approximately 65% of the total industry capital expenditure and 25% of the value added. It is concentrated primarily in East Asia (Taiwan, South Korea and Japan) and mainland China.

[5] SIA and BCG report *“Government Incentives and US Competitiveness in Semiconductor Manufacturing*”, September 2020

Assembly, packaging & testing

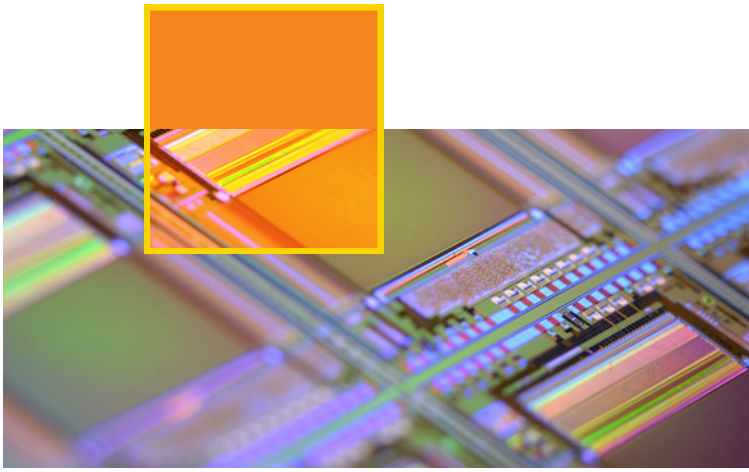
(back end manufacturing)
3% of industry R&D
13% of industry capex
6% of the value added

This stage involves converting the silicon wafers produced by the fabs into finished chips that are ready to be assembled into electronic devices. Firms involved at this stage first slice silicon wafers into individual chips. Chips are then packaged into protective frames and encased in a resin shell. Chips are further rigorously tested before being shipped to electronic device manufacturers.

The back-end stage of the supply chain still requires significant investments in specialized facilities. Firms specializing in assembly, packaging and testing typically invest over 15% of their annual revenues in facilities and equipment. Although it is relatively less capital-intensive and employs more labor than the front-end fabrication stage, new innovations in advanced packaging are changing this dynamic. Overall, this activity accounts for 13% of the total industry capital expenditure and contributed 6% of the total value added by the industry in 2019. It is concentrated primarily in Taiwan and mainland China, with new facilities also being built recently in Southeast Asia (Malaysia, Vietnam, and the Philippines).

A highly specialized support ecosystem

Companies focused on semiconductor design and production activities are supported by an upstream ecosystem of specialized suppliers.



Electronic design automation (EDA)

3% of industry R&D
<1% of industry capex
4% of the value added

At the design stage, electronic design automation (EDA) companies provide sophisticated software and services to support designing semiconductors, including outsourced design of specialized application specific integrated circuits (ASICs). With billions of transistors in a single chip, state-of-the-art EDA tools are indispensable to design competitive modern semiconductors.

Core IP suppliers license reusable components designs – commonly called “IP blocks” or “IPs” – with a defined interface and functionality to design firms to incorporate into their chip layouts. These also include foundation physical IPs associated with each manufacturing process nodes, as well as many interface IPs. EDA and core IP vendors invest heavily in R&D – about 30 to 40% of their revenues – and accounted for approximately 4% of the value added of the industry in 2019.

Wafer processing and testing equipment

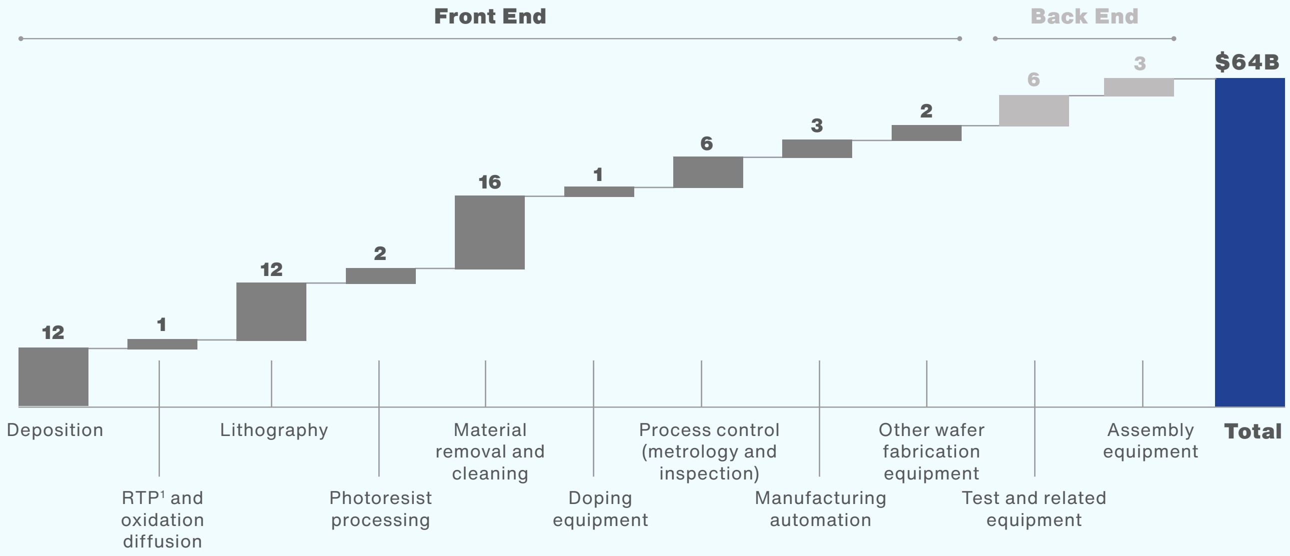
9% of industry R&D
3% of industry capex
11% of the value added

Semiconductor manufacturing uses more than 50 different types of sophisticated wafer processing and testing equipment provided by specialist vendors for each step in the fabrication process. (Exhibit 8).

EXHIBIT 8

Semiconductor production involves more than 50 types of sophisticated specialized equipment

Breakdown of market size of semiconductor manufacturing equipment by major families, 2019 ($ Billion)



# Equipment types											
12	4	6	1	12	4	16	2	2+	3	5	50+

1. Rapid thermal processing
Sources: Gartner

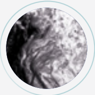
Lithography tools represent one of the largest capital expenditures for fabrication players and determine how advanced of a chip a fab can produce. Advanced lithography equipment, specifically those that harness Extreme Ultra-Violet (EUV) technology are required to manufacture chips at 7 nanometers and below. A single EUV machine can cost $150 million.

Metrology and inspection equipment is also critical for the management of the semiconductor manufacturing process. Because the process involves hundreds of steps over one to two months, if any defects occur early in the process, all the work undertaken in the subsequent time-consuming steps will be wasted. Strict metrology and inspection processes using specialized equipment are therefore established at critical points of the semiconductor manufacturing process to ensure that a certain yield can be confirmed and maintained.

Modern fabs also have advanced automation and process control systems for direct equipment control, automated material transportation and real-time lot dispatching, with many of the newest facilities almost entirely automated.

Semiconductor manufacturing equipment also incorporates many subsystems and components with specific functionality, such as optical or vacuum subsystems, gas and fluid management, thermal management or wafer handling. These subsystems are provided by hundreds of specialized suppliers.

Developing and fabricating such advanced, high-precision manufacturing equipment also requires large investments in R&D. Semiconductor manufacturing equipment companies typically invest 10 to 15% of their revenues in R&D. Overall semiconductor equipment manufacturers suppliers accounted for 9% of the R&D and 11% of the value added of the industry in 2019.



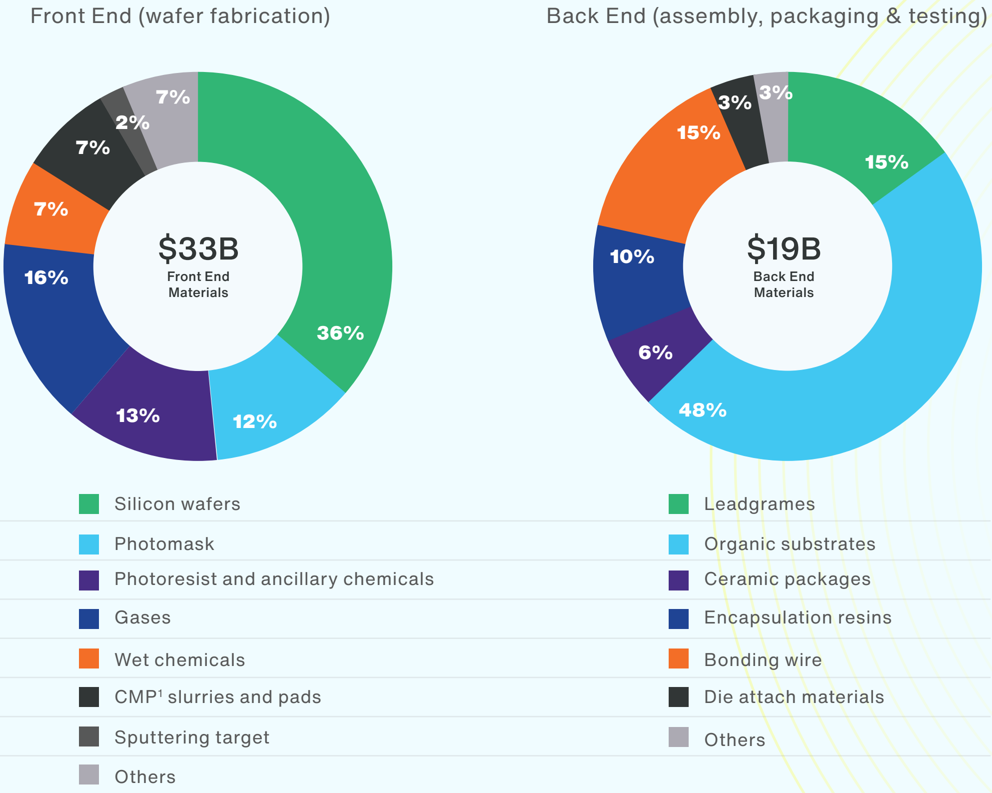
Materials
1% of industry R&D
6% of industry capex
5% of the value added

Finally, firms involved in semiconductor manufacturing also rely on specialized suppliers of materials. Semiconductor manufacturing uses as many as 300 different inputs, many of which also require advanced technology to produce. For example, the polysilicon employed to make the silicon ingot that is subsequently sliced into wafers is required to have a purity level that is 1,000 times higher than the level required for solar energy panels, and is provided primarily by just four companies, with a combined global market share above 90%. Exhibit 9 shows the breakdown of the global sales of semiconductor manufacturing materials in 2019 across the key families used in front-end and back-end manufacturing.

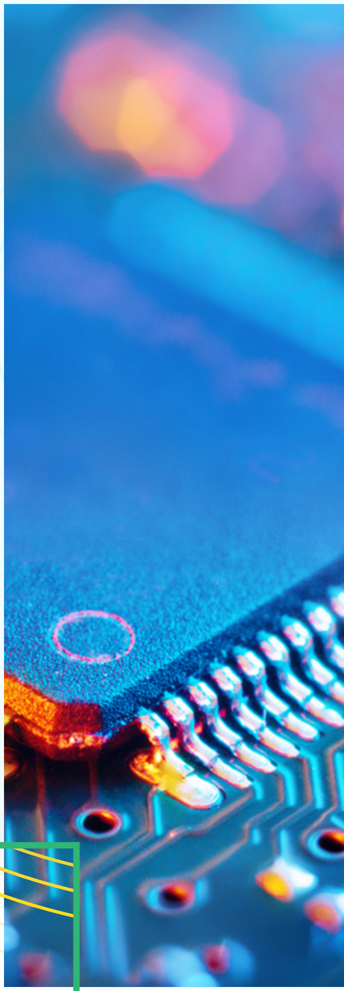
EXHIBIT 9

Semiconductor production uses hundreds of unique materials and specialty chemicals

Breakdown of market size of semiconductor manufacturing materials, 2019 (% of $ Billion)



1. Chemical-mechanichal planarization
Sources: BCG analysis based on data from SEMI, IHS and HSBC



The main front-end materials include:

Polysilicon: is a metallurgical grade silicon in ultra-refined purity levels, suitable for use in semiconductor wafer production.

Silicon wafers: Polysilicon is melted, formed into single crystal ingots which are then sliced into wafers, cleaned, polished, and oxidized in preparation for circuit imprinting within fabrication facilities.

Photomask: A plate covered with patterns used in the lithography process. The patterns consist of opaque and clear areas that prevent or allow light through.

Photoresist: A special material that undergoes a chemical reaction upon exposure to light. Silicon wafers are covered with a photoresist layer, which is imprinted with the patterns contained in the photomask during the lithography process.

Wet processing chemicals: Used in the etching and cleaning steps of the semiconductor manufacturing process, and include solvents, acids, etchants, strippers and other products.

Gases: Used to protect wafers from atmospheric exposure. Other gases are used in the semiconductor manufacturing process as dopants, dry etchants, and in chemical vapor deposition (CVD).

Chemical Mechanical Planarization (CMP) slurries: Materials used for polishing the surface of the wafer after the film deposition step to provide a flat surface.

Back-end materials include leadframes, organic substrates, ceramic packages, encapsulation resins, bonding wires and die-attach materials. They typically have relatively lower technical barriers to produce compared to the wafer fabrication materials described above.

Production of these highly specialized materials is done in large plants, which also require high investments. Annual capital expenditure by the leading global suppliers of silicon wafers, photoresistors or gases typically ranges between 13 and 20% of their revenues. Overall, materials suppliers contributed 6% of the total capital expenditure and accounted for 5% of the value added of the industry in 2019.

The unique simultaneous high R&D and high capital intensity of semiconductors

Semiconductors are very complex products to design and manufacture. As a result, the semiconductor industry presents both high R&D and high capital intensity. Overall, we estimate that in 2019 the industry invested about $90 billion in R&D and $110 billion in capital expenditure globally across all the activities in the value chain. These two figures combined represent almost 50% of the $419 billion in global semiconductor sales in the same year.

As Exhibit 4 above showed, while 65% of the total industry R&D investment (excluding pre-competitive research) is made in the design layer of the value chain, there is also significant R&D activity in EDA and core IP, semiconductor equipment and wafer manufacturing. Similarly, 65% of the total industry capital expenditure is incurred for wafer manufacturing, but assembly and test, materials and even design also require significant investments in advanced facilities and equipment.

Considering the investments made by firms across the entire global value chain, no other industry has the same high level of intensity in both R&D (22% of annual final chip revenues, ahead of pharmaceuticals) and capital expenditure (26% of final chip revenues, ahead of utilities). (See Exhibit 10.) This extremely high level of investment intensity creates the need for large global scale and specialization.

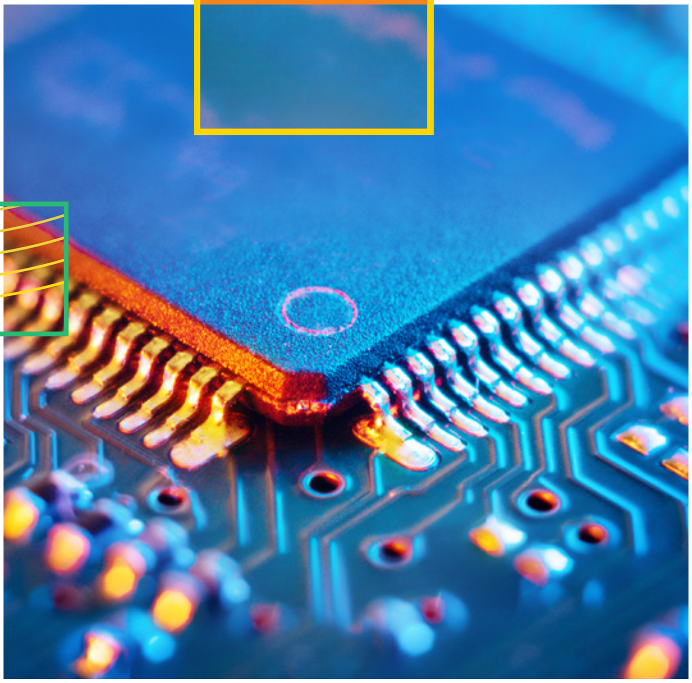
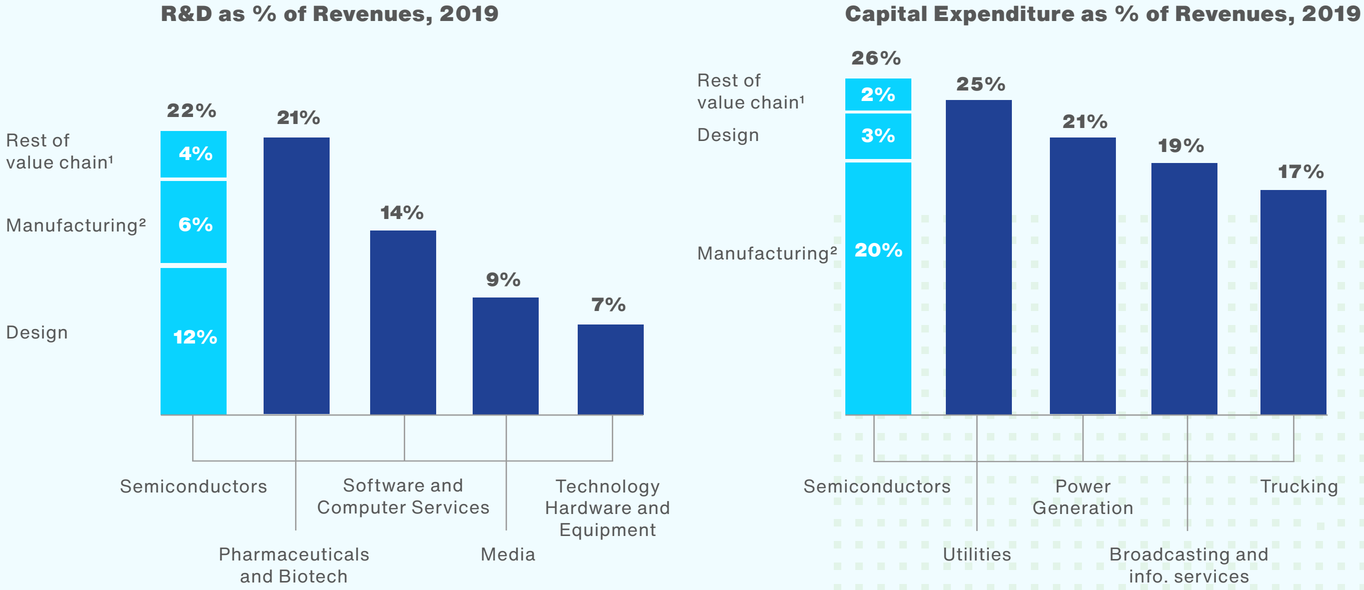


EXHIBIT 10

The semiconductor industry ranks high simultaneously in both R&D and capital intensity



1. Includes EDA and Core IP, Equipment and Materials
2. Includes Wafer Fabrication and Assembly, Packaging & Testing
Sources: BCG analysis based on Capital IQ data

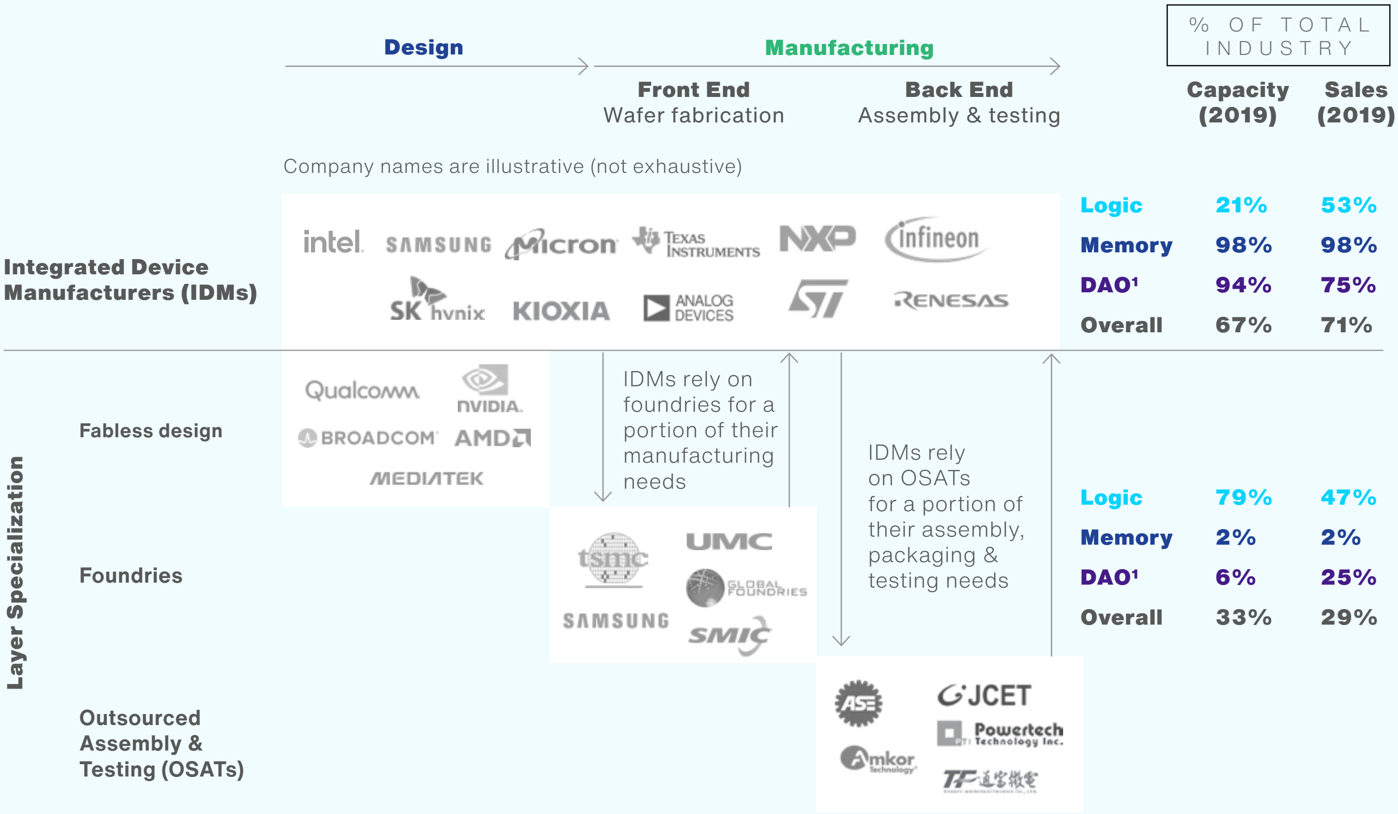
Semiconductor business models

Since the inception of the semiconductor industry in the 1960s, its structure has evolved from its original form of only vertically integrated firms doing all stages of production. The dramatic increase in technology complexity and need for scale to afford massive investments to keep the pace of innovation in both design (in the form of R&D) and manufacturing (in the form of capital expenditure) favored the emergence of specialized players.

Today semiconductor companies may focus on one layer of the supply chain or integrate vertically across several layers. No company or even entire nation is vertically integrated across all. There are four types of semiconductor companies, depending on their level of integration and business model (Exhibit 11): integrated device manufacturers (IDMs), fabless design firms, foundries and outsourced assembly and test companies (OSATs).

EXHIBIT 11

Technology complexity and need for scale have led to emergence of business models focused on a specific layer of the value chain



IDMs rely on foundries for a portion of their manufacturing needs

IDMs rely on OSATs for a portion of their assembly, packaging & testing needs

% OF TOTAL INDUSTRY

Integrated Device Manufacturers (IDMs)

	Capacity (2019)	Sales (2019)
Logic	21%	53%
Memory	98%	98%
DAO¹	94%	75%
Overall	67%	71%

Layer Specialization

	Capacity (2019)	Sales (2019)
Logic	79%	47%
Memory	2%	2%
DAO¹	6%	25%
Overall	33%	29%

1. Discrete, analog and optoelectronics and sensors

Sources: BCG analysis with data from SIA WSTS, Gartner

IDMs
Segment economics
(% of annual revenue, 2016-2019)
- Gross Margin: 52%
- R&D: 14%
- Capex: 20%
- Operating Cash Flow: 17%

IDMs are vertically integrated across multiple parts of the value chain, performing design; fabrication; and assembly, packaging and test activities in-house. In practice, some IDMs have hybrid “fab-lite” models where they outsource some of their production and assembly.

In the early decades of the industry, the IDM model was predominant, but the rapidly increasing size of the investments in both R&D and capital expenditure created the simultaneous need for both scale and specialization, which led to the emergence of the fabless-foundry model.

Currently the IDM model is more common for firms focused on memory and DAO products, which are largely general-purpose components and more scalable. IDMs accounted for approximately 70% of the global semiconductor sales in 2019.

Fabless
Segment economics
(% of annual revenue, 2016-2019)
- Gross Margin: 50%
- R&D: 20%
- Capex: 4%
- Operating Cash Flow: 20%

Fabless firms choose to focus on design and outsource fabrication as well as assembly, packaging, and testing. Fabless firms typically outsource fabrication to pure-play foundries and OSATs. The fabless model has grown along with the demand for semiconductors since the 1990s as the pace of innovation made it increasingly difficult for many firms to manage both the capital intensity of manufacturing and the high levels of R&D spending for design. As technical difficulty and upfront investment soared with the migration to smaller manufacturing nodes, total semiconductor sales accounted for by fabless firms increased from less than 10% in 2000 to almost 30% in 2019.

Logic chips are fundamentally the realm of fabless firms with the notable exception of Intel and, more recently and at a lower scale, Samsung. This dynamic is due to the pace at which the market demands improved power and performance capabilities in order to support the quick cycles of smartphones and emerging leading-edge applications in AI and high-performance computing.

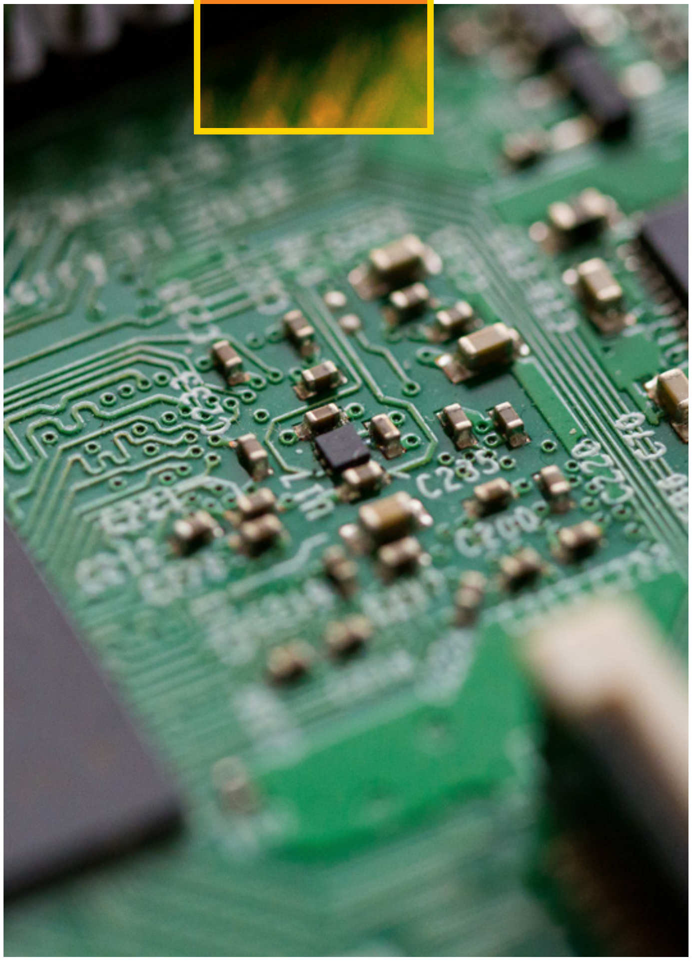
Foundries

Segment economics
(% of annual revenue, 2016-2019)

- Gross Margin: 40%
- R&D: 9%
- Capex: 34%
- Operating Cash Flow: 15%

Foundries address the fabrication needs of fabless firms and IDMs alike, as most IDMs do not have sufficient installed manufacturing capacity in-house to cover all their needs. This business model enables foundries to diversify the risk associated with the large upfront capital expenditure required to build modern fabs across a larger customer footprint of design firms and IDMs. Most foundries are focused purely on manufacturing for third parties, although some IDMs with strong manufacturing capabilities may also choose to make chips for others in addition to their own.

Leaving memory aside, foundries have added 60% of the incremental capacity in the industry for DAO and logic products during the past five years. Currently foundries account for 35% of the total industry manufacturing capacity, or 50% if memory is excluded. Their share rises to 78% in advanced (14 nanometers or below) and trailing nodes (20 to 60 nanometers) using the more advanced 12”/300mm wafer size. Furthermore, the only two companies that can currently manufacture at the leading 5 nanometer node are foundries.



OSATs

Segment economics
(% of annual revenue, 2016-2019)

- Gross Margin: 17%
- R&D: 4%
- Capex: 16%
- Operating Cash Flow: 2%

OSATs provide assembly, packaging and test services under contract to both IDMs and fabless companies. This part of the supply chain was first offshored by some US IDMs starting back in the 1960s because of its lower capital intensity and the need for lower-skilled labor. The fabless-foundry model then also led to the emergence of specialized OSAT companies.

A need for massive global scale

The economics described above, together with the deep expertise in the complex technology required to produce semiconductors, create natural barriers to entry across the core activities in the supply chain, leading to a relatively concentrated supplier base in each activity.

In manufacturing, the sheer size of the upfront investment required to build new capacity acts as a major barrier. As an illustration, the aggregated annual capital expenditure of the top 5 foundries between 2015 and 2019 amounted to approximately $75 billion, or an average of $3 billion per firm per year, equivalent to more than 35% of their annual revenues.

While semiconductor design does not require large amounts of capital expenditure, its high R&D intensity also creates significant scale advantages and acts as a barrier to entry. For example, the top 5 fabless firms invested $68 billion in R&D in the 5 years between 2015 and 2019, or an average of $2.8 billion per firm per year, equivalent to 22% of their revenues.

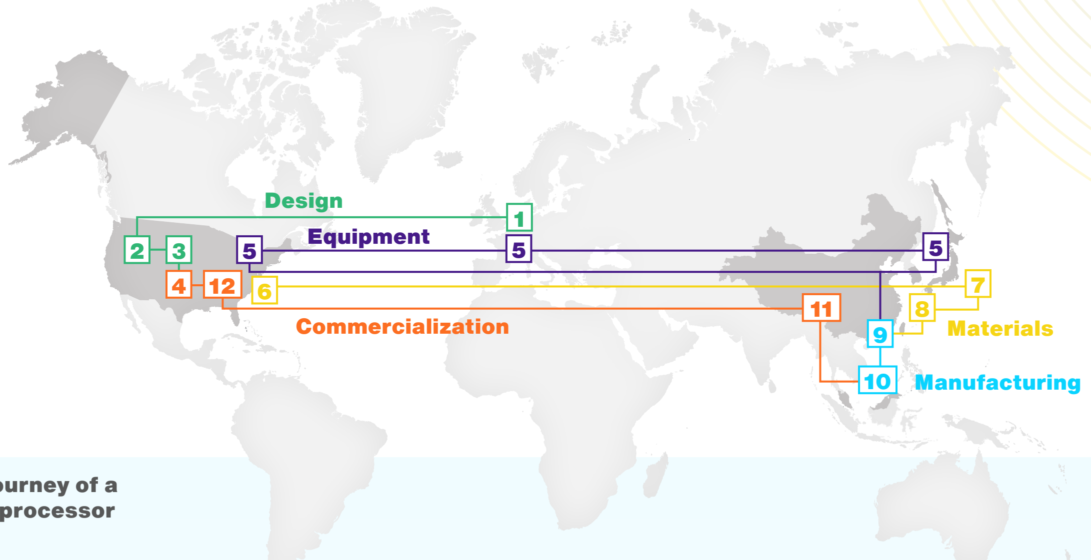
Achieving a satisfactory return on these massive investments is possible only for firms with very large scale. This is the reason why, across the different activities in the semiconductor supply chain, the top 3 players globally generally account for between 50% and 90% of their respective segment revenues.

A Global Integrated Structure Based on Geographic Specialization

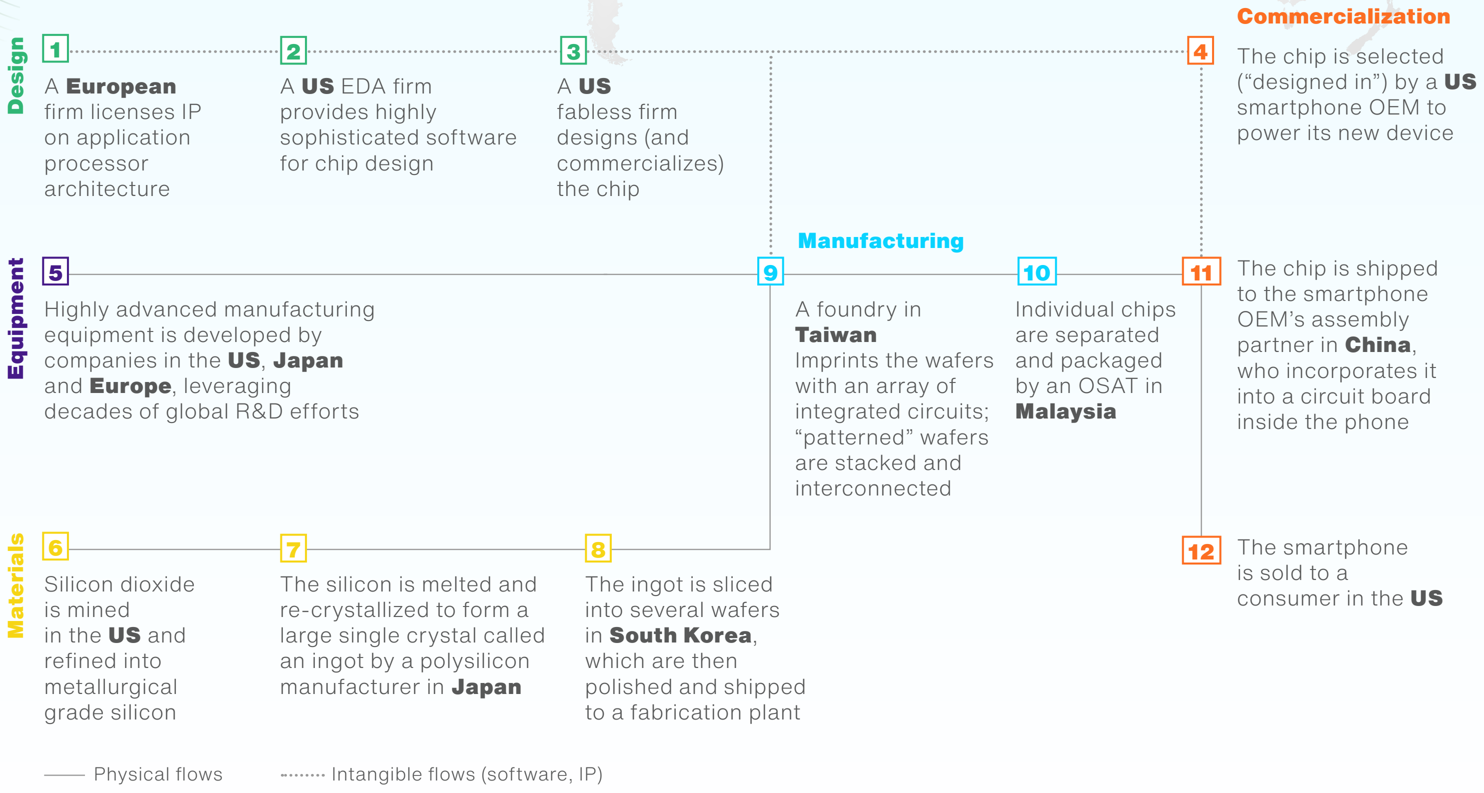
The semiconductor supply chain is truly global: six major regions (US, South Korea, Japan, mainland China, Taiwan and Europe) each contribute 8% or more to the total value added by the semiconductor industry in 2019. As Exhibit 12 shows, the typical journey of a semiconductor involves most if not all of these geographic areas at different stages during the design and manufacturing process.

EXHIBIT 12

The semiconductor value chain is truly global and relies on the specialized capabilities of different geographic areas



Illustrative: The global journey of a smartphone application processor



1 A **European** firm licenses IP on application processor architecture

2 A **US** EDA firm provides highly sophisticated software for chip design

3 A **US** fabless firm designs (and commercializes) the chip

4 The chip is selected (“designed in”) by a **US** smartphone OEM to power its new device

5 Highly advanced manufacturing equipment is developed by companies in the **US**, **Japan** and **Europe**, leveraging decades of global R&D efforts

6 Silicon dioxide is mined in the **US** and refined into metallurgical grade silicon

7 The silicon is melted and re-crystallized to form a large single crystal called an ingot by a polysilicon manufacturer in **Japan**

8 The ingot is sliced into several wafers in **South Korea**, which are then polished and shipped to a fabrication plant

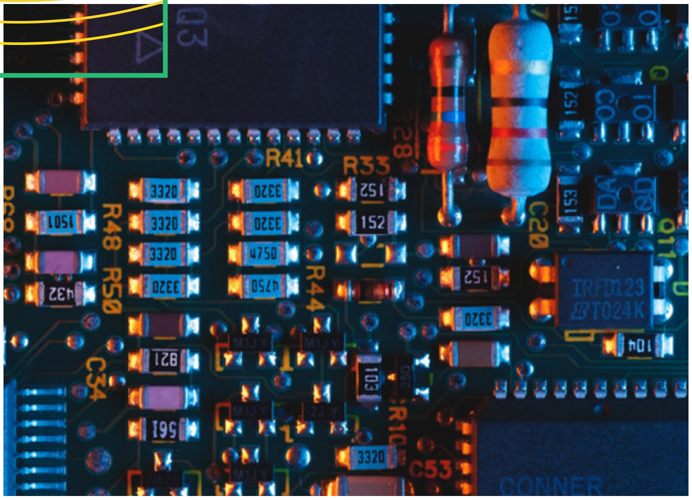
9 A foundry in **Taiwan** Imprints the wafers with an array of integrated circuits; “patterned” wafers are stacked and interconnected

10 Individual chips are separated and packaged by an OSAT in **Malaysia**

11 The chip is shipped to the smartphone OEM’s assembly partner in **China**, who incorporates it into a circuit board inside the phone

12 The smartphone is sold to a consumer in the **US**

—— Physical flows ········ Intangible flows (software, IP)



To some extent, this highly distributed structure of the semiconductor supply chain follows the global geographic footprint of the electronics industry that the semiconductor industry serves. Proximity to the leading firms that develop these devices may be important for semiconductor design companies. Different regions have strength in certain types of end electronics devices or applications:

- The US is the global leader in the design of electronic devices. As shown on Exhibit 2, US consumer electronics, information technology, automotive and industrial firms source 35% of the semiconductors used in the world, with particular strength in advanced chips for PCs and data centers.
- Greater China (including Taiwan) is the largest manufacturing global hub for electronic devices. Together, local original equipment manufacturers (OEMs) and contract manufacturers that assemble devices designed by other companies based elsewhere are responsible for more than 60% of the world's production of consumer electronics, smartphones, and PCs. Proximity to these companies that are the ultimate destination of the components to be assembled into devices may be important for semiconductor manufacturing companies.
- Europe is the global leader in automotive and industrial automation equipment; Japan is strong in these two sectors and also consumer electronics; South Korea is a global force in smartphones and other consumer electronics, too.

While proximity to customers is clearly a significant driver of the global, interdependent structure of the semiconductor supply chain, there are three additional key factors:

- **Global R&D networks.** International collaboration has allowed multinational companies, universities and institutions to collaborate and pool resources. Together, they undertake pre-commercial research to pursue scientific breakthroughs that lead to major leaps in semiconductor technology.
- **Geographic specialization.** Regions are focused primarily on different activities within the semiconductor supply chain. This regional division of tasks has been driven by comparative advantage developed over the past decades of the industry's history.
- **Trade liberalization.** Global trade policies enable participants in the semiconductor industry to move goods, equipment, capital, IP and talent across borders, effectively supporting the geographic specialization across the semiconductor supply chain.

Need for deep technical know how and massive scale have created a highly specialized global supply chain

We now look at these three factors in more detail.

Global R&D networks

A significant portion of the R&D investment by the semiconductor industry is in fundamental research into science breakthroughs, invested many years ahead of a potential commercial application. Semiconductor firms and institutions such as universities and government-funded advanced science labs typically collaborate on this pre-competitive research to share the costs of research and avoid duplication of efforts.

China and the US are the top two countries in scientific publications related to semiconductors filed in the past 10 years[6]. However, our analysis of scientific publications shows that fundamental semiconductor research often involves collaboration across borders:

- 36% of the scientific publications related to semiconductors published by Chinese institutions were co-authored with institutions from other countries – in fact, the US was the largest research partner for Chinese institutions.
- In the case of publications from US institutions, 60% were co-authored with institutions from other countries, with China as the largest partner followed by Germany and South Korea.

---

[6] According to BCG analysis of ~150,000 scientific publications related to semiconductors published between January 2010 and December 2020, based on data sourced from Derwent Innovation and LexisNexis PatentSight

In addition, the semiconductor industry has created or is a core contributor to a number of organizations that bring together global companies, universities and research institutions to support international collaboration in R&D, such as IMEC, CEA-Leti and A*STAR (see sidebar).

Select examples of leading global semiconductor research institutions

- **The Interuniversity MicroElectronics Centre (IMEC)** was established in Belgium in 1984 as a non-profit organization. Today, IMEC is an international innovation hub in nanoelectronics, semiconductors and other digital technologies. It has research labs in seven countries across Europe, Asia and North America, with more than 4,000 researchers from about 100 different nationalities. IMEC has a vast global network of over 600 partners, including governments and governmental agencies, universities and corporations.
- **CEA-Leti** was established in France in 1967 as a non-profit research branch of the French Alternative Energies and Atomic Energy Commission (CEA). Today it is one of the world's largest research institutes for applied research in microelectronics and nanotechnology, helping companies to bridge the gap between basic research and manufacturing. CEA-Leti has 1,900 researchers (about 40% of whom are foreign nationals) and 250 active industrial partners including multiple leading global semiconductor firms.
- **Singapore's Agency for Science, Technology and Research (A*STAR)** was established in 1991 to foster R&D that is aligned to areas of competitive advantage and national needs for Singapore, currently spanning four technology domains set by the nation's five-year R&D plan. It collaborates with leading global companies through joint research programs hosted in its advanced R&D lab facilities. In semiconductors A*STAR is pioneering joint research with industry partners in areas such as beyond 5G radio frequency (RF) technologies, power electronics, advanced packaging, microelectromechanical systems (MEMS), and artificial intelligence hardware with in-memory computing.
- **Semiconductor Research Corporation (SRC)** was established in the US in 1982. It brings together over 100 universities, 20 leading semiconductor companies and 3 government agencies with the objective of accelerating pre-competitive academic research in microelectronics. SRC pools industry and government funds to create larger and more effective research programs in breakthrough semiconductor technology areas that are undertaken by its network of over 2,000 SRC-sponsored university researchers.

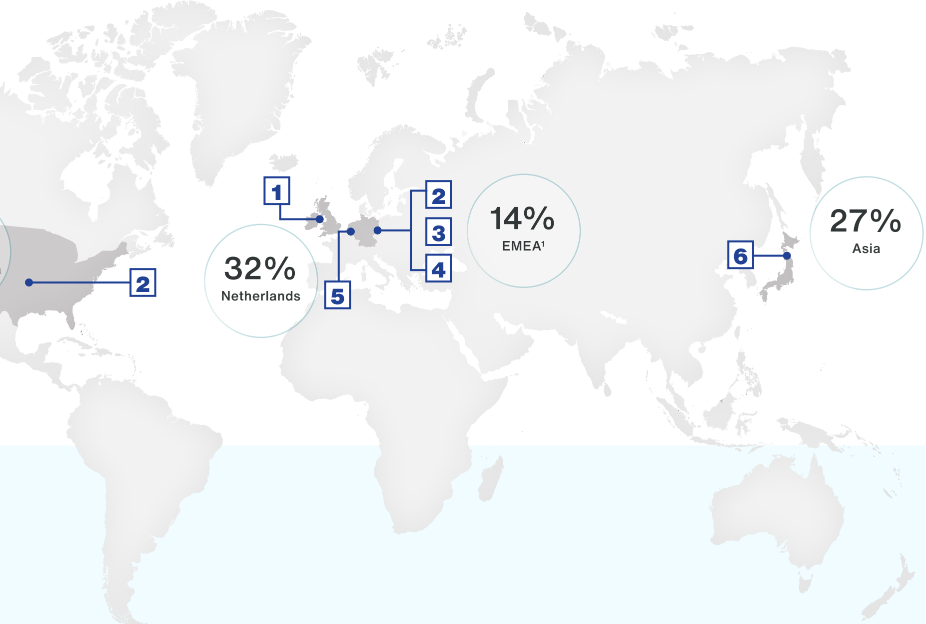
In fact, some of the most critical recent advancements in semiconductor technology were the result of several decades of global R&D collaboration. Fin field-effect transistor (FinFET) technology enabled manufacturing at 22 nanometers and is the dominant transistor design for today's leading-edge chips at 5 nanometers. While the US pioneered the development of the FinFET technology and is the source of 48% of the related patents, the rest of the world also contributed heavily to applied R&D leading to commercialization. Specifically, Taiwan, which hosts several of the world's leading foundries, has contributed 20% of the FinFET patents.

In the case of EUV, the technology that underpins the equipment utilized to manufacture semiconductors on the 7- and 5-nanometer nodes and below, its development started in the 1980s with fundamental research done in the US and Japan on the use of soft X-rays, leading to the first demonstration of the technology in 1986. In the 1990s and early 2000s, NTT in Japan, Bell Labs and Lawrence Livermore National Laboratory in the US and the University of Twente in the Netherlands further pushed the research in this technology. ASML, a company based in the Netherlands, then sought to further develop and commercialize EUV in partnership with institutions like IMEC and corporations including Intel (headquartered in the US), Samsung (South Korea), and TSMC (Taiwan). Together, ASML and its global partners funded R&D during the pre-commercial stage of the technology, and subsequently ASML invested $8 billion to put the technology in production in modern fabs beginning in 2018.

In addition to the global collaboration in the development of the underlying technology, EUV also relies on a global supply chain: as Exhibit 13 shows, today the EUV lithography equipment developed by ASML contains about 100,000 parts provided by over 5,000 suppliers spread across the globe.

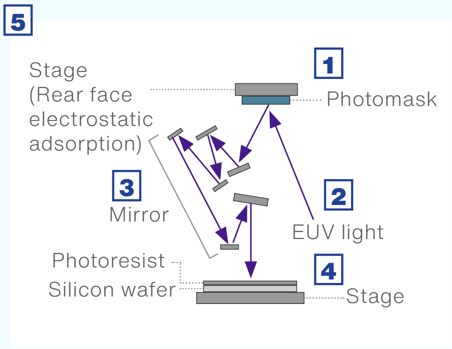
EXHIBIT 13

EUV equipment integrates components from a global network of more than 5,000 specialized suppliers



○ **Share of total number of suppliers**

Sample EUV key components



1 UK

Vacuum system
Edwards: Keeps system in vacuum to minimize EUV absorption by air

2 US

EUV light source
Cymer² : System uses CO2 laser to vaporize 50k molten tin droplets to generate EUV light (13.5nm wavelength)
Note: ¾ of light source costs are from Europe

Other tool parts
Mechatronics/electronics, ultra low thermal expansion glass for optics

2 Germany

Laser & power source
Trumpf: CO2 laser supported by 1MW power supply and advanced cooling system

3 Germany

Optical system
Zeiss³ : Precise mirrors that project light on wafer for patterning

4 Germany

Wafer chuck
Berliner Glass²: Electrostatic chuck used to clamp wafer during litho process

5 Netherlands

Vessel
VDL: Modular housing for EUV light source

6 Japan

Other tool parts
Structural ceramics, photoresist and photomask

1. EMEA excluding Netherlands 2. Subsidiary of ASML 3. ASML owns minority interest
Source: ASML; expert interview; BCG research and analysis

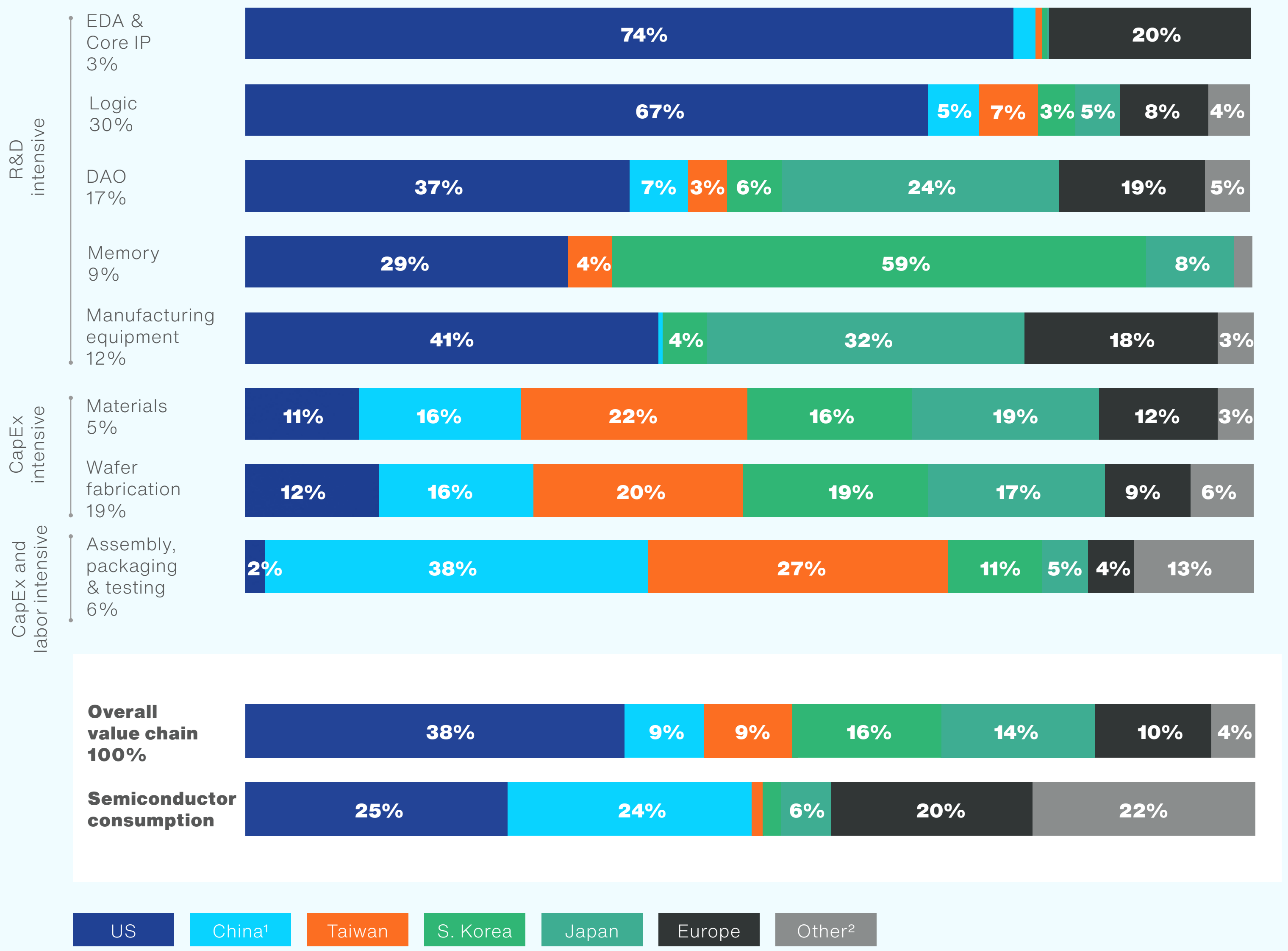
Geographic specialization

As mentioned above, six major regions have a significant participation in the total global output of the semiconductor industry. But each region plays a different role in the global semiconductor supply chain: broadly speaking, the US leads in the activities that are most intensive in R&D: EDA and core IP, chip design and manufacturing equipment. Raw materials and manufacturing (both wafer fabrication as well as assembly, packaging and testing), which are more capital intensive, are largely concentrated in Asia (Exhibit 14).

EXHIBIT 14

Regions specialize in different activities of the value chain: US leads in R&D intensive activities; Asia leads in manufacturing

Semiconductor industry value added by activity and region, 2019 (%)



1. Mainland China 2. Other includes Israel, Singapore and the rest of the world
Notes on regional breakdown: EDA, design, manufacturing equipment and raw materials based on company revenues and company headquarters location. Wafer fabrication and Assembly & testing based on installed capacity and geographic location of the facilities
Sources: BCG analysis with data from company financials, Capital IQ, Gartner, SEMI, IHS Markit

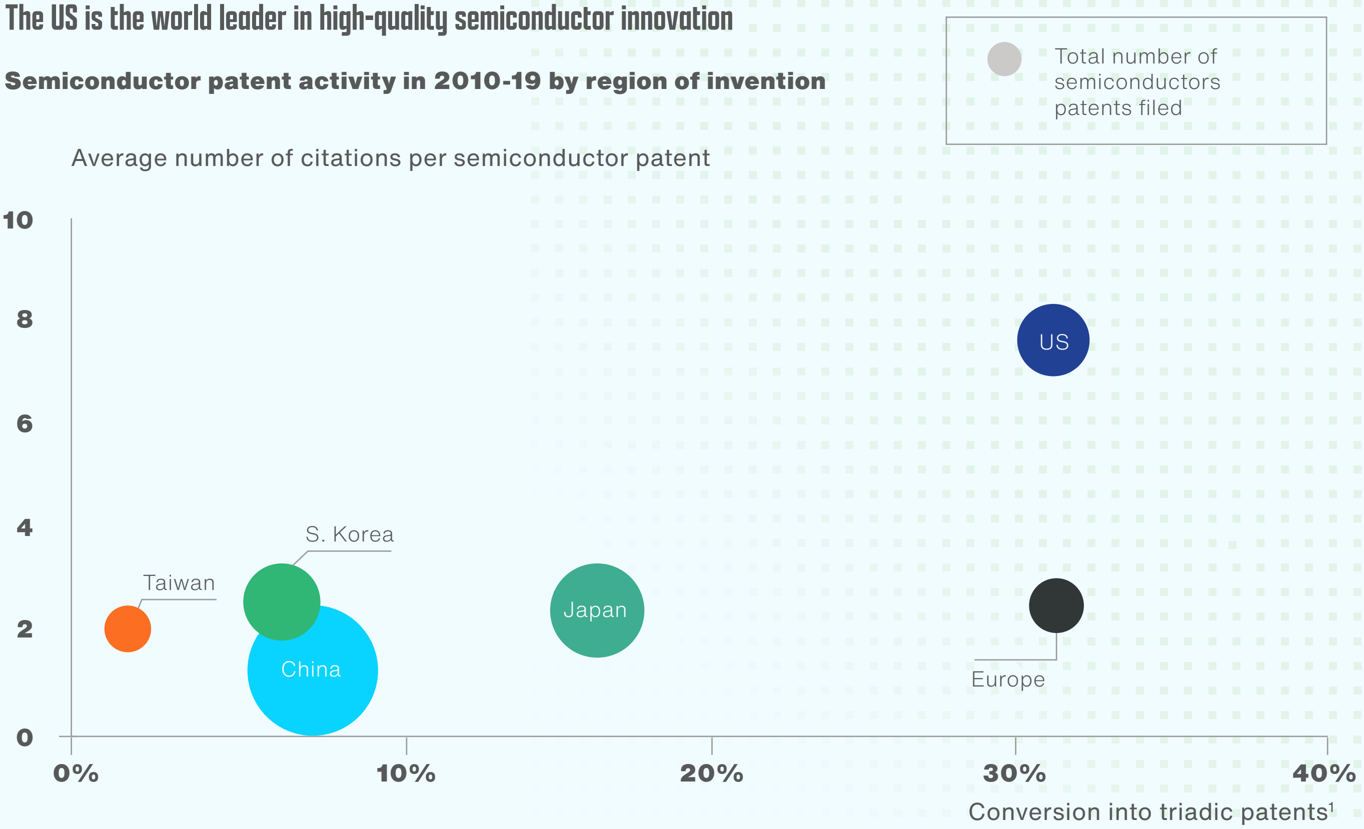
Specialization based on comparative advantage underpins this differentiated regional focus across the semiconductor supply chain.

Despite the stagnation in government support for basic research in the last decades described ealier, the US is the global leader in semiconductor R&D. It is home to some of the world's most prominent clusters of technical universities and semiconductor companies, which has resulted in a virtuous cycle of education, research, entrepreneurship, and access to capital to fuel innovation. While China has been investing aggressively in semiconductor R&D and currently files the largest total number of semiconductor academic research papers and patents annually, the US is still the source of the most relevant innovation in the industry: together with Europe, it has the highest conversion from patents filed into triadic patents – typically regarded as a marker of high-quality innovation with global commercial potential – in semiconductors, and the average number of citations per US semiconductor patent is between three and six times higher than for patents from any other country in the world (Exhibit 15).

EXHIBIT 15

The US is the world leader in high-quality semiconductor innovation

Semiconductor patent activity in 2010-19 by region of invention

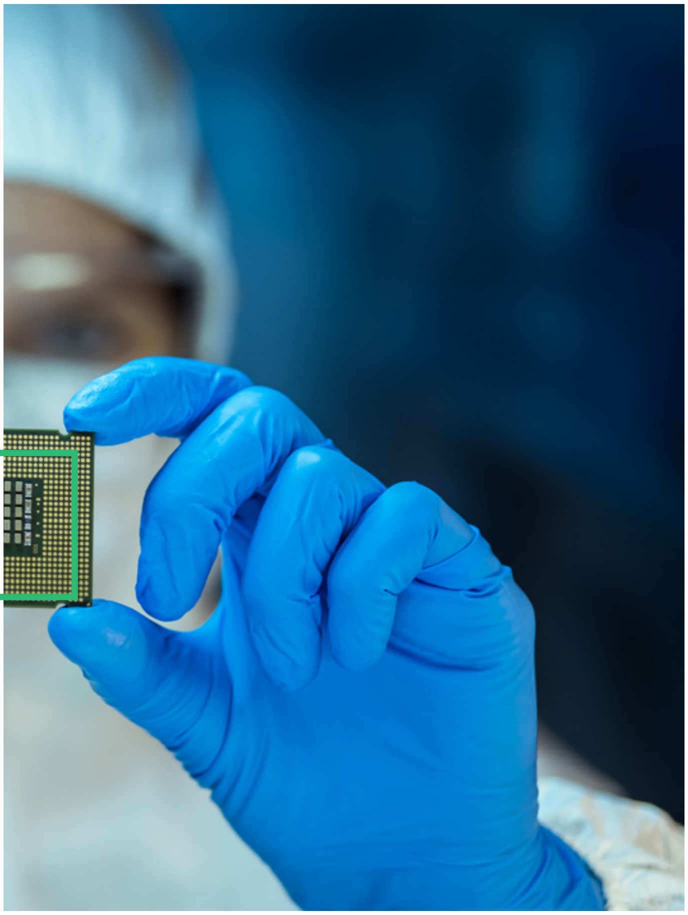


1. Triadic patents are those filed in the world's largest markets (US, EU and Japan) and are typically considered to cover higher-value inventions. Conversion ratio calculated as total number of semiconductor triadic patents per OECD database in 2010-16 divided by the total numbers of semiconductor patents filed in home patent office and at least one other major patent office (US, Europe, Japan, China or South Korea) in 1999-2015
Source: OECD Patent Database, Derwent Innovation, LexisNexis PatentSight, BCG Center for Growth & Innovation Analytics (GIA)

In **chip design**, where US companies – including both fabless firms and IDMs—have a combined share of almost 50% of global semiconductor sales, the critical success factors are access to highly skilled engineering talent and a thriving innovation ecosystem, especially leading universities. 10 of the top 20 semiconductor design companies (including both fabless and IDMs), as well as 4 of the top 5 EDA and core IP companies by 2019 revenue, are headquartered in the US. US design companies invest more heavily in R&D than their peers: according to their financial statements, on average US companies involved in semiconductor design spent 18% of their 2019 revenues on R&D.

---

7. Triadic patents are a series of corresponding patents filed at the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO) and the Japan Patent Office (JPO), for the same invention, by the same applicant or inventor



A look into workforce location also highlights the US leadership position in semiconductor design: about 50% of the engineers employed by the top global semiconductor companies involved in design are located in the US[8]. This figure includes engineers from both US and non-US firms. In parallel, design companies increasingly rely on access to global engineering talent pools, particularly in India, where we estimate that 20% of the world's semiconductor design engineers sit today.

Foreign talent contributes decisively to the US leadership in semiconductor innovation. Looking back in history, some of the most fundamental semiconductor technology breakthroughs that are still used in most modern semiconductors – such as the metal–oxide–semiconductor field-effect (MOSFET) transistor and the complementary metal–oxide–semiconductor (CMOS) fabrication process -- were developed by immigrants to the US. At present about 40% of the high-skilled semiconductor workers in the US were born abroad, according to a recent study by the Center for Security and Emerging Technology (CSET)[9]. This is linked to the large pipeline of global talent coming from US technical universities: international students comprise around two-thirds of graduate students in electrical engineering and computer science, and more than 80% stay in the country after completing their degrees.

In **fabrication**, which is extremely capital intensive, the availability of attractive investment conditions – particularly government incentives – and access to robust infrastructure (power and water supply, transportation and logistics) and a skilled manufacturing workforce at competitive rates have traditionally been the key success factors. As shown in our prior report focused on semiconductor manufacturing economics, government incentives may account for up to 30-40% of the 10-year total cost of ownership (TCO) of a new state-of-the-art fab, which is estimated to amount to $10-15 billion for an advanced analog fab and $30-40 billion for advanced logic or memory.

US leads in R&D intensive activities, supported by its "talent magnet", Asia leads in capital intensive activities, supported by government incentives

The East Asia region (including Japan, South Korea and Taiwan) and mainland China currently concentrate about 75% of the world's total semiconductor manufacturing capacity– including all the leading edge capacity at 7 nanometers and below currently in operation – and under current market conditions its share is expected to continue rising over the next decade. According to our analysis summarized in Exhibit 16, the TCO of a new fab located in the US is approximately 25-50% higher than in Asia, and 40-70% of that difference is attributable directly to government incentives, which are currently much lower in the US than in alternative locations.

---

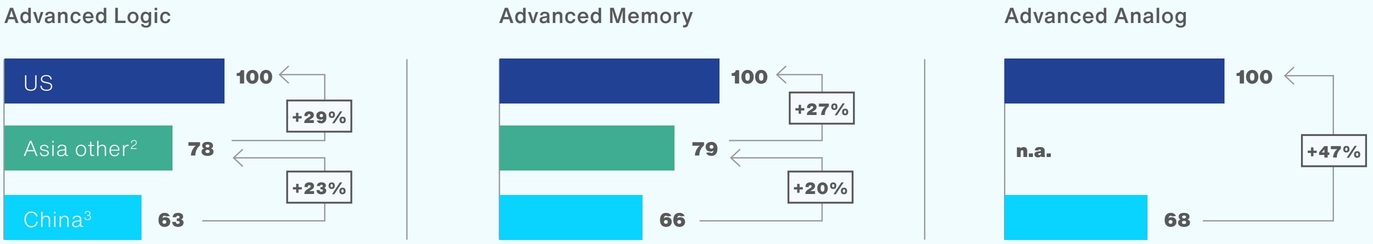
[8] The total number of design related positions has been estimated based on publicly available profiles on Linkedin for the top 10 fabless and IDM players. The number may be underestimated for certain regions such as China due to the availability of public data.

[9] Center for Security and Emerging Technology (CSET), *The Chipmakers. U.S. Strengths and Priorities for the High-End Semiconductor Workforce*, September 2020

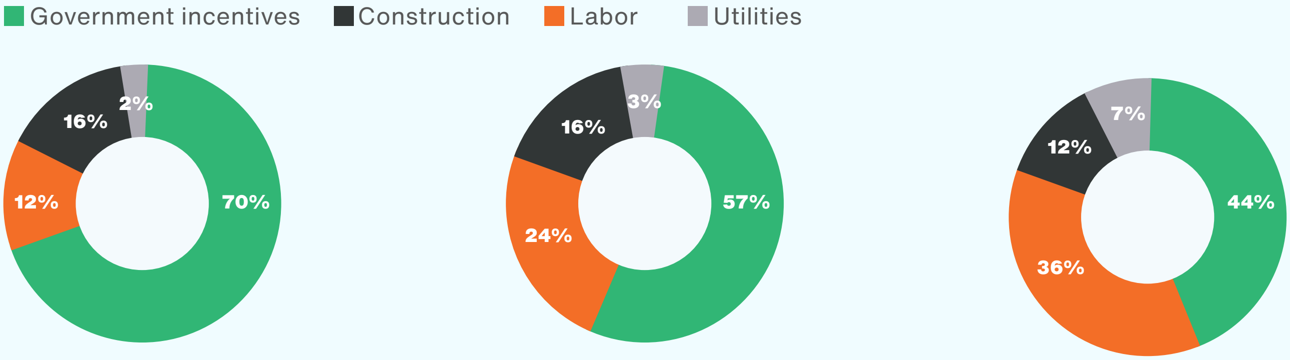
EXHIBIT 16

Manufacturing economics are significantly more favorable in Asia, with government incentives driving most of the cost advantage

Estimated 10-year Total Cost of Ownership (TCO[1]) of reference fabs by location (US indexed to 100)



What drives the higher TCO of US-based fabs
% of incremental TCO for US-based fabs



1. TCO includes capital expenditure (upfront land, construction and equipment) + 10 years of operating expenses (labor, utilities, materials, taxes) 2. Refers to Taiwan and South Korea for logic, South Korea and Singapore for memory 3. With technology sharing agreements that give access to additional incentives such as equipment lease back with advantageous terms
Source: BCG analysis

In particular, Taiwan has been investing in the development of its domestic semiconductor manufacturing industry since 1974, when the government selected semiconductors as a key focus industry to expand the economy beyond agriculture. Policies pursued by the government included both direct support in the form of setting up R&D labs and industrial parks and providing incentives for the construction of new fabs such as generous tax-credits that could cover as much as 35% of their capital expenses and 13% of their equipment purchases, as well as indirect incentives such as reform of the financial sector and capital markets to facilitate access to funding. While several of the incentive programs were reduced after 2009-2010, we estimate that Taiwan still provides incentives for new fabs worth 25-30% of their overall total cost of ownership over a 10-year period. This is in line with other Asian locations such as South Korea and Singapore, but currently well below mainland China. In contrast, the incentives to new fab construction currently available in the US and Europe are estimated to reach just 10-15% of the total cost of ownership[10].

---

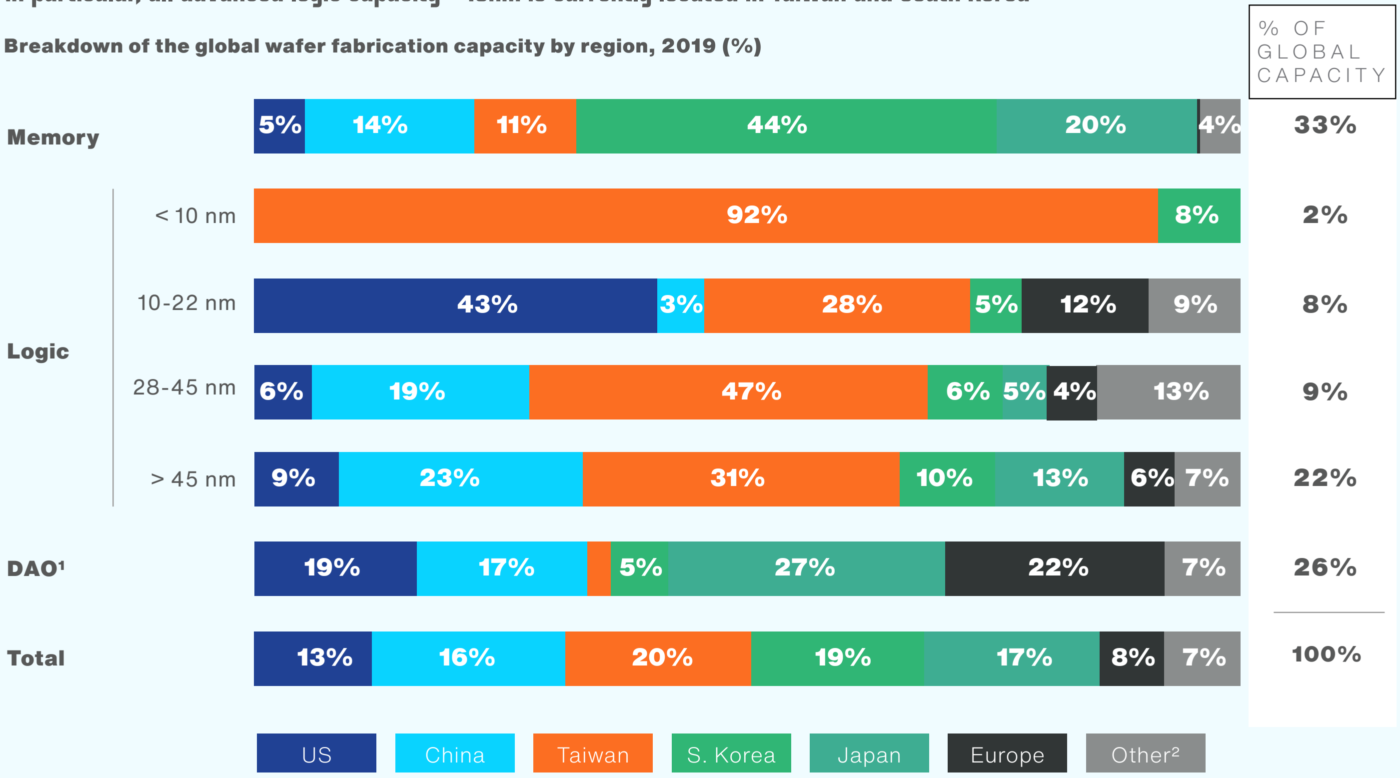
[10] For further discussion of government incentives for semiconductor manufacturing, see our prior report "Government Incentives and US Competitiveness in Semiconductor Manufacturing", September 2020.

In the late 1980s and 1990s Taiwanese firms pioneered the foundry model, specializing in manufacturing the chips designed by firms from other regions. Today Taiwan is home to 2 of the 5 largest foundries globally and hosts 20% of the total global capacity. Along with Intel (US) and Samsung (South Korea), TSMC is one of three firms that can produce logic chips in advanced nodes (10 nanometers or below), which are required for compute-intensive devices such as data center/AI servers, PCs, and smartphones. In fact, almost all of the world's capacity in the leading nodes (5 and 7 nanometers) is located in Taiwan (Exhibit 17).

EXHIBIT 17

East Asia + China concentrate about 75% of the wafer fabrication capacity; in particular, all advanced logic capacity < 10nm is currently located in Taiwan and South Korea

Breakdown of the global wafer fabrication capacity by region, 2019 (%)



1. Discretes, analog and optoelectronics and sensors
2. Other includes Israel, Singapore and the rest of the world

Sources: BCG analysis with data from SEMI fab database

In comparison, **assembly, packaging and testing** is much less capital intensive. While the annual capital expenditure of foundry companies is typically around 35% of their revenues, for leading firms specializing in outsourced semiconductor assembly and testing (OSATs) capital expenditure typically runs at less than half of that level, at approximately 15% of their revenues. Given the lower capital intensity, the cost of labor is a key competitive factor for OSAT firms.

According to data from the Conference Board[11], the average manufacturing wages for skilled labor in mainland China, Taiwan and Southeast Asian countries such as Singapore and Malaysia are up to 80% below US levels. Today 9 of the 10 largest OSAT firms by revenue are headquartered in mainland China, Taiwan and Singapore. In terms of capacity location, mainland China and Taiwan account for more than 60% of the world's assembly, packaging and testing capacity. Recently OSAT firms have also started to diversify their own global footprint, building new capacity in other locations with low labor costs such as Malaysia. However, with the increasing level of technology innovation in the field of advanced packaging, labor cost may become less of a decisive factor going forward.

[11] The Conference Board: *International Comparisons of Hourly Compensation Costs in Manufacturing*, 2018

Trade liberalization

The geographic specialization described above means that firms focused on a particular layer of the semiconductor supply chain need to interact and collaborate with other firms upstream or downstream in the chain that are typically located in other countries. Furthermore, given that semiconductors are used in all types of electronics products, ultimately the semiconductor components need to be shipped to where the manufacturing of end devices occurs.

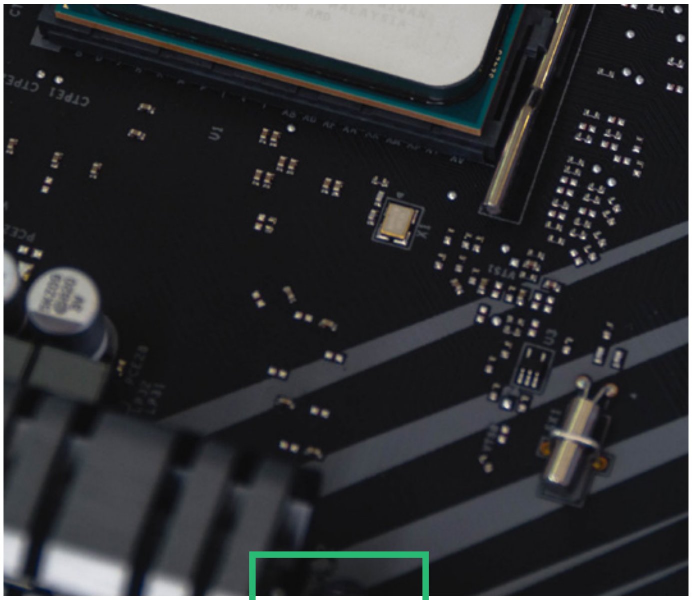
For example, as shown earlier in this report, while we estimate that US-based electronic device makers are responsible for sourcing 33% of the total global semiconductor demand, in many cases their devices are actually manufactured outside the US. Less than 20% of total semiconductors sales are actually shipped to the US to be integrated into a product. Instead China is the destination of approximately 35% of the shipments of semiconductors as many of the world's leading electronic device makers have their products assembled there—an activity further downstream in the electronics supply chain that is more labor intensive and with lower valued add.

Semiconductors are the world's fourth most traded product

These two factors create the need to move materials, tools/equipment, products and IP across borders, which has been enabled by international trade agreements that eliminated tariffs and trade barriers for semiconductor products and reinforced the protection of intellectual property. Integrated circuits are one of the products subject to the lowest tariffs in global trade[12]. In particular, the World Trade Organization's Information Technology Agreement (ITA) effective since 1997 and further expanded in 2015, has been instrumental for the strong growth in international trade of semiconductor related products. Our analysis shows that the trade in semiconductor-related goods included in the original 1997 ITA grew at 10.5% CAGR over a 20-year period, outpacing the rest of the semiconductor products not covered in the agreement by 3 points of annual growth and generating a 20% increase in the value of global semiconductor-related trade.

Indeed, the global nature of the supply chain and the interdependencies between countries are well illustrated by the magnitude and composition of the semiconductor trade flows (Exhibit 18). In 2019, global semiconductor trade reached $1.7 trillion in trade value[13]. This is more than four times the value of 2019 global semiconductor sales, indicating the large magnitude of cross-border transactions involved in the development and manufacturing of semiconductors. In fact, semiconductors are the world's 4th most traded product, only after crude oil, motor vehicles and parts and refined oil.

According to our analysis, more than 120 different countries (over 60% of the countries in the world) were involved as an exporter or importer of semiconductor products, signifying the scope and reach that the semiconductor industry has in the world. And even though China's share of semiconductor design or manufacturing is still relatively low, the country's preeminent position in the manufacturing and assembly of electronic devices has allowed it to emerge as a central hub for semiconductor trade.



---

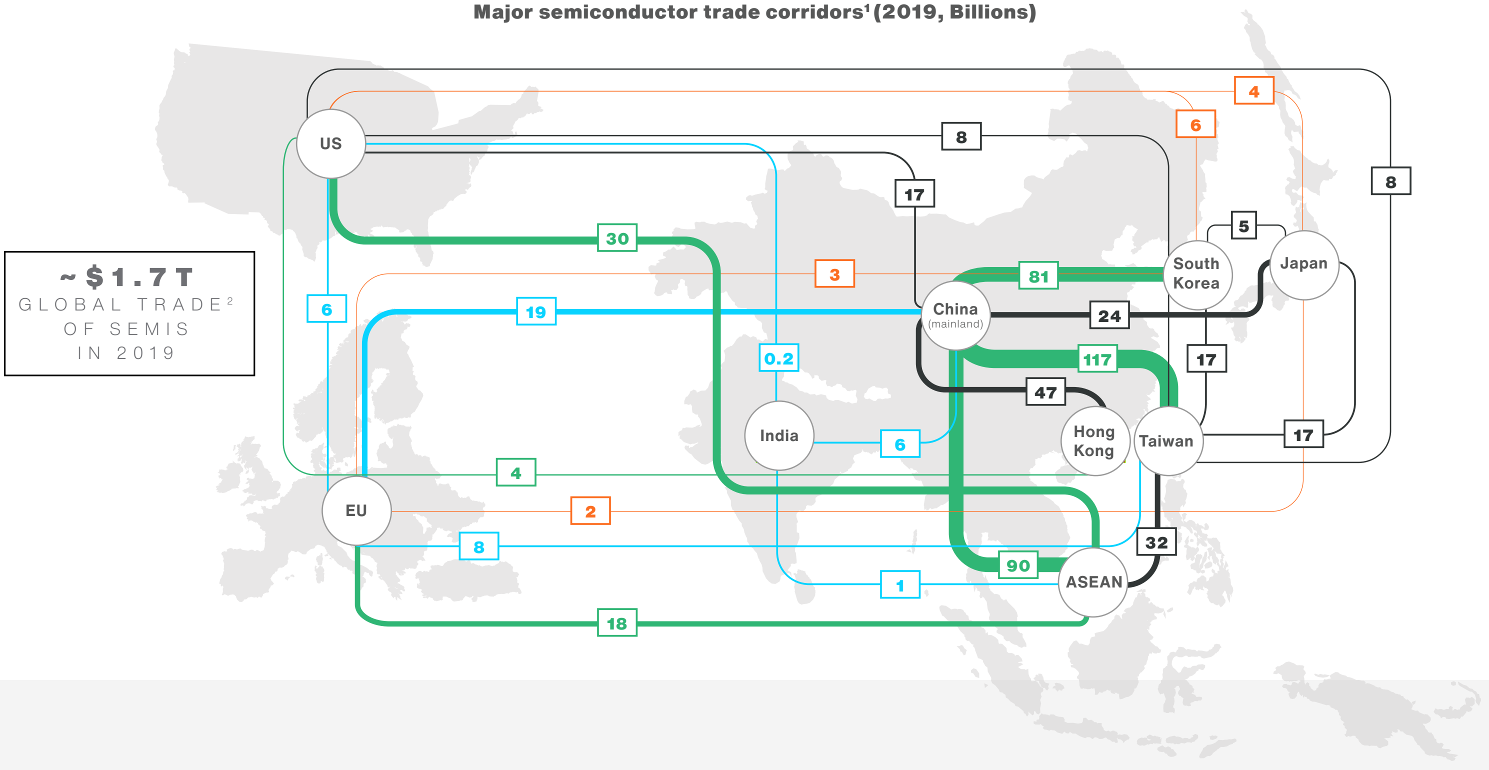
[12] According to data compiled by the Observatory of Economic Complexity (OEC), integrated circuits (HS code 8542) ranks in the position 1236 out of 1259 products in terms of tariffs applied across the world.

[13] Includes exports + imports classified in HS codes 8542 (integrated circuits) and 8541 (semiconductor discrete devices), minus HS 854140 (Photosensitive, photovoltaic, LED semiconductor devices). Does not include semiconductor equipment or materials

EXHIBIT 18

A large web of global trade flows supports the geographic specialization in the semiconductor value chain

Major semiconductor trade corridors[1] (2019, Billions)



Width of arrow represents size of trade flows in 2019

Color of arrow represents 2014-2019 CAGR in trade

<0% | 0-5% | 5-10% | >10%

1. HS codes 8541, 8542, minus HS 854140, excludes semiconductor equipment
2. Includes both exports and imports. Note: Significant disparities in reported data by each country. Importer data used where possible;

Source: IHS Global Trade Atlas, UN Comtrade; BCG analysis

Implication of this global structure: mutual interdependence

The global structure of the semiconductor supply chain, with geographic specialization across layers, means that companies interact and collaborate across borders, in relationships of mutual dependency.

For example, while the US is the clear global leader in several layers of the supply chain (EDA and core IP, design, manufacturing equipment) characterized by high R&D intensity, and has a share of global semiconductor sales (45-50%) well above its share of the global consumption of electronic devices by end users (25%), it still depends on other countries for many activities, mainly in semiconductor manufacturing: materials, wafer fabrication, assembly, packaging and testing services, and even some key advanced equipment required for manufacturing in the leading nodes such as EUV lithography.

This mutual dependency resulting from specialization based on comparative advantage brings tremendous benefits for the semiconductor industry, and ultimately for the makers of electronics devices that rely on continuous improvements in the performance and cost of semiconductors to drive advances in digital services.

As an illustration of the benefits from this global structure, we consider a hypothetical scenario where US semiconductor firms had to have all their products manufactured onshore. Since US companies accounted for 49% of the global semiconductor sales in 2019, this means that in this hypothetical scenario the US would have 49% of the global semiconductor manufacturing capacity onshore, instead of the current 12%. Without adequate government incentives, these fabs located in the US would have comparatively higher operating costs (labor, electricity, as well as the incremental annual depreciation due to higher upfront capital expenditure including the difference in government incentives across regions) than the existing ones located in South Korea, Taiwan or mainland China. Using the fab economics model from our September 2020 report, we estimate that in this hypothetical scenario the cost of production for US semiconductor companies would increase by about 15%. In turn, this would undermine the competitiveness of US-based semiconductor firms and reduce their ability to sustain the current R&D investment levels. Given the global leadership of US firms in chip design, it could ultimately slow down innovation and ultimately result in higher costs for electronic device makers across the world.

Emerging Risks in the New Global Context

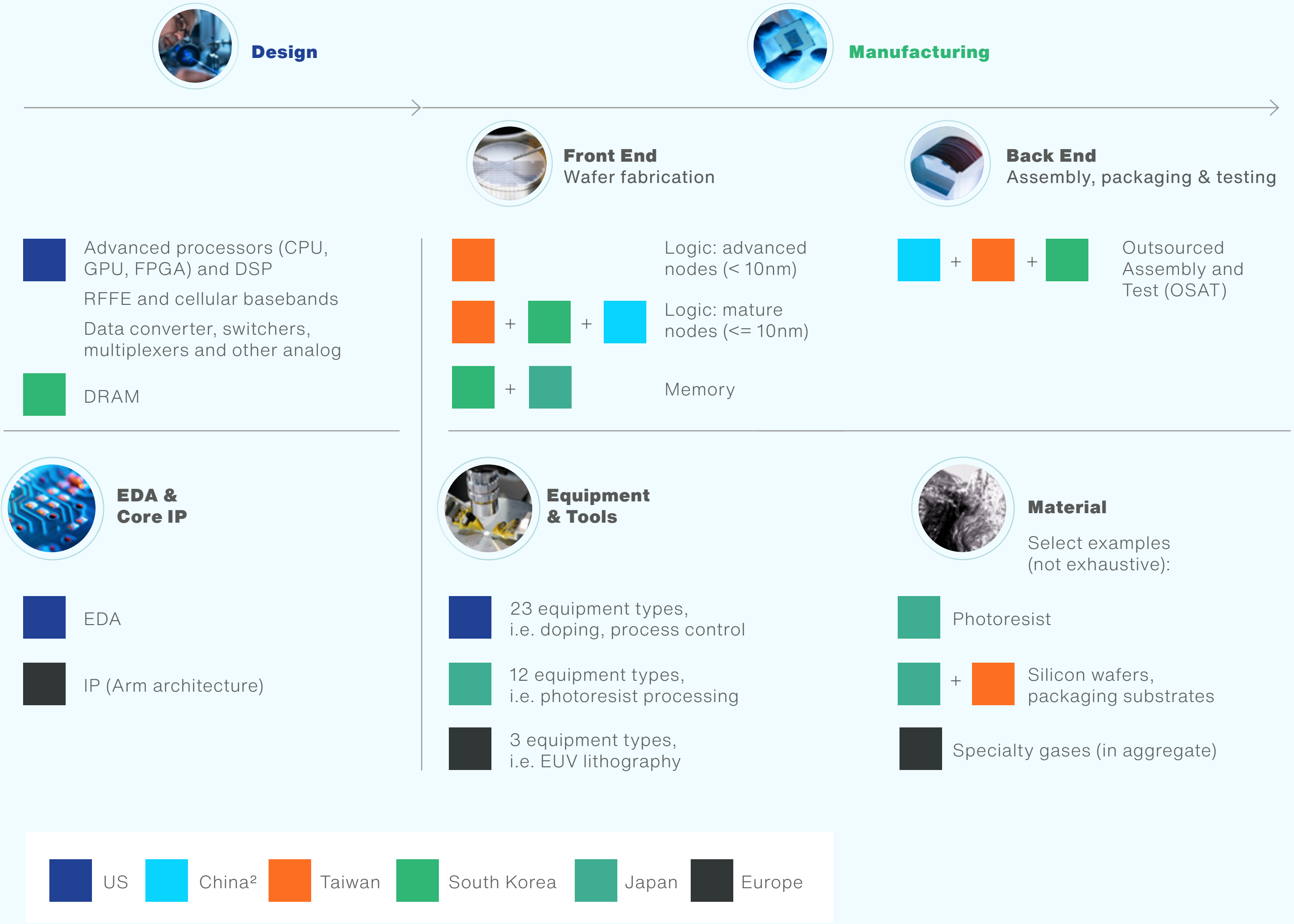
The global structure of the semiconductor supply chain, developed over the course of the past three decades, has served the industry well. Ultimately, it has enabled the explosion in innovation and end user adoption of information technology, which has benefited consumers and businesses immensely. However, in the last few years several new factors have emerged that could put the successful continuation of this global model at risk.

Over the last three decades, the benefits of geographic specialization based on comparative advantage have resulted in the emergence of a more concentrated and interdependent global semiconductor supply chain. While not exhaustive, our analysis shows that there are more than 50 points across the overall supply chain where a single region accounts for 65% or more of the total global supply (Exhibit 19).

EXHIBIT 19

Multiple points of high geographical concentration across the current semiconductor value chain

Value chain activities where one single region accounts for ~65% or more of global share[1]



Design

🟦 Advanced processors (CPU, GPU, FPGA) and DSP
RFFE and cellular basebands
Data converter, switchers, multiplexers and other analog

🟩 DRAM

Manufacturing

Front End
Wafer fabrication

🟧 Logic: advanced nodes (< 10nm)

🟧 + 🟩 + 🔷 Logic: mature nodes (<= 10nm)

🟩 + 🟢 Memory

Back End
Assembly, packaging & testing

🔷 + 🟧 + 🟩 Outsourced Assembly and Test (OSAT)

EDA & Core IP

🟦 EDA

⬛ IP (Arm architecture)

Equipment & Tools

🟦 23 equipment types, i.e. doping, process control

🟢 12 equipment types, i.e. photoresist processing

⬛ 3 equipment types, i.e. EUV lithography

Material

Select examples (not exhaustive):

🟢 Photoresist

🟢 + 🟧 Silicon wafers, packaging substrates

⬛ Specialty gases (in aggregate)

🟦 US 🔷 China[2] 🟧 Taiwan 🟩 South Korea 🟢 Japan ⬛ Europe

1. For Design, EDA & Core IP, Equipment & Tools and Raw Materials: global share measured as % of revenues, based on company headquarter location. For Manufacturing (both Front End and Back End) measured as % of installed capacity, based on location of the facility 2. Mainland China
Sources: BCG analysis with data from Gartner, SEMI, UBS; SPEEDA

While geographic specialization has served the industry well, the observed high degree of geographic concentration in certain activities also creates two types of vulnerabilities:

- Single points of failure due to high geographic concentration of some activities that could result in large-scale supply interruptions
- Geopolitical tensions that may impair global access to suppliers or customers

Single points of failure that create risk of large-scale supply interruptions

Excessive geographic concentration in manufacturing exposes the industry to single points of failure which may be disrupted by natural disasters, infrastructure failures, cyberattacks or geopolitical frictions – from tariffs and export controls to even supply blockage resulting from broad embargoes or armed conflicts.

Numerous examples of such disruptions, albeit contained in scale and duration to date, can be found in the last thirty years (see sidebar).

Manufacturing is clearly a major point of concern for the resilience of the global semiconductor supply chain. Currently almost 75% of the global installed capacity is concentrated in East Asia (Japan, South Korea and Taiwan) and mainland China, a region significantly exposed to high seismic activity and geopolitical tensions. The number is even higher for advanced technologies: currently 100% of the global capacity in the leading 7- and 5-nanometer nodes is currently in East Asia.

> High geographic concentration of manufacturing capacity has emerged as a concern for the resilience of the semiconductor supply chain

In particular, as shown before in Exhibit 17, Taiwan has 40% of the world's logic chip production capacity and leads in the most advanced nodes at 10 nanometers or below that are required to manufacture chips such as application processors, CPUs, GPUs and FPGAs for smartphones, PCs, data center servers, and autonomous vehicles. In an extreme hypothetical scenario of complete disruption for one year, Taiwanese foundries would lose their current cumulative $42 billion in revenues,

Historical examples of disruption to semiconductor supply

- The impact of the explosion of a Sumitomo Chemical factory in Japan in 1993 is often cited to illustrate the magnitude of this risk. It impacted 60% of the global supply of epoxy resin, and spot prices for DRAM memory chips in the US market spiked from an average of $30/megabyte to around $80/megabyte.
- A strong earthquake in the center of Taiwan in September 1999 caused a six-day shutdown of the Hsinchu Science Park due to power outages. As a result, memory-chip prices tripled and shares of electronics companies around the world tanked, with IBM, Hewlett Packard, Intel, and Xerox, all part of the Fortune 100 at the time, losing 18 to 40% of their value in the month after the earthquake.
- In 2011 a major earthquake struck Japan, followed by a tsunami and nuclear power-plant melt down. 25% of the global production of silicon wafers and 75% of the global supply of hydrogen peroxide was affected by the disaster. Several fabs were shut down for several months.
- In 2019, geopolitical tensions between Japan and South Korea rose sharply. Japan imposed export controls on semiconductor materials to Korea, impacting approximately $7 billion in semiconductor exports per month.
- In December 2020, a power outage affected a memory fab located in Taiwan for just one hour, impacting 10% of global DRAM supply.
- Two fires at a package substrate plant in Taiwan in October 2020 and February 2021 aggravated the global capacity shortage for assembly, packaging and testing services, which was already experiencing difficulties to meet the surge in semiconductor demand in the last few months of 2020.
- Widespread power failures following a polar vortex in Texas, and a fire in a Renesas fab in Japan in early 2021 further exacerbated the global chip supply shortage, especially for the automotive market.

---

14. Based on the estimated share of device applications that are supplied by chips produced by Taiwanese fabs, including PC/laptop/tablet, servers, smartphone, automotive electronics, and industrial cameras.

but that could also cause a $490 billion drop in revenue, or 12 times more negative impact, for electronic device makers across different application markets[14]. The global electronics supply chain would come to a halt, creating significant global economic disruptions. If such hypothetical complete disruption were to be become permanent, it could take a minimum of three years and $350 billion of investment in what would be an unprecedented effort to build enough capacity in the rest of the world to replace the Taiwanese foundries.

A high degree of geographic concentration of supply also exists in some materials, such as silicon wafers, photoresist, some chemicals such as packaging substrates, or specialty gases. While each specialty material accounts for only a tiny portion of the industry's total value added, semiconductors cannot be fabricated without them.

As an example, C4F6 is a critical process gas used to make 3D NAND memory and some advanced logic chips. It is essential for the etching process during chip fabrication, allowing etching to be completed 30% faster than the nearest alternative. Furthermore, once a manufacturing plant is calibrated to use C4F6, it cannot be substituted. Sales of C4F6 were approximately $250 million in 2019, with the top three suppliers located in Japan (40% of global supply), Russia (25%), and South Korea (23%). If any of these top three producers were severely disrupted, the loss of $60-100 million in C4F6 supplies, could lead to about $10 to $18 billion of lost revenue for NAND alone downstream in the semiconductor chain – almost 175 times higher than the direct impact. If such disruption in a portion of C4F6 supply were to become permanent, NAND production levels would potentially be constrained for 2-3 years until alternative locations could introduce new capacity ready for mass production.

Geopolitical tensions that may impair global access to suppliers or customers

While not exposing the industry to the risk of immediate halt of manufacturing activity leading to component shortages for electronic device makers, geographic concentration of the ownership of the leading global suppliers – measured in terms of company headquarters location as a proxy for where the technology is actually developed – also exists in other points of the semiconductor value supply chain (see Exhibit 19 above).

- **In semiconductor manufacturing equipment**, US firms collectively account for more than 50% share of the global market in 5 of the major manufacturing process equipment categories (deposition tool, dry/wet etch and cleaning, doping equipment, process control, and testers). Likewise, Japan has over a 90% share of the photoresist processing market, vital equipment to the lithography process. In addition, ASML – a European company – has practically a 100% global market share in the EUV lithography machines essential to manufacture on advanced nodes below 7 nanometers.
- US-headquartered firms collectively account for more than 90% share in **advanced logic products** such as CPUs, GPUs or FPGAs that power PCs, data center servers, AI analytics and automotive ADAS systems – although manufacturing of these products is largely done in Asian foundries.
- Likewise, three US-based firms – of which one now has a European parent company - have a combined 85% share in **EDA software** tools essential to design semiconductors
- In the **core IP** layer, Arm - a company headquartered in the UK, but with R&D operations in multiple locations including the US – licenses the architecture and processor core designs that currently run practically every smartphone and an increasing portion of the embedded computing systems used in IoT applications in Consumer Electronics, Industrial and Automotive.

Under normal market conditions this may not present immediate supply issues. In some cases potential substitutes may exist in other countries, and these activities are typically easier to scale than wafer manufacturing. However, they could also be subject to disruptions in scenarios of trade or geopolitical conflict that introduce restrictions to access to suppliers or technology originated in certain countries.

Overall, geopolitical tensions have been rising globally in the last 10 years: the index measuring global geopolitical risk is back at the levels of the Gulf War in 1990-1991. Ongoing geopolitical tensions in key semiconductor trade corridors in Asia and between the US and China present a rising risk to the industry supply chain.

Japan-South Korea tensions

The origin of the dispute is South Korea’s continued push for restitution for Japan’s World War II era transgressions – claims that Japan contends were settled in a 1965 treaty. Following a series of Korean court rulings against Japanese companies on this matter, Japan imposed restrictions on exports to South Korea in July 2019.

Among the more than 1,000 products potentially exposed to Japanese export controls, three chemicals key to semiconductor manufacturing were of particular concern to South Korea: hydrogen fluoride (Japanese vendors account for 70% of the global supply), fluorinated polyimide and photoresists (in both cases Japanese vendors account for 90% of the global supply). While the value of South Korean imports of these three inputs from Japan is relatively small—only about $400 million annually—South Korea exports $80 billion worth of semiconductors each year that depend on those chemicals, 250 times greater than the revenue loss of Japan’s chemical manufacturers.

Given South Korea’s preeminent position in the global semiconductor supply chain — it is the second largest semiconductor manufacturer in the world, with a 44% share of the global market in memory – the impact from this conflict could go beyond the semiconductor industry and disrupt the entire global electronics supply chain downstream.

While it appears that tensions cooled down somewhat during 2020, and Japan has been approving export license requests for the three chemicals in question, the situation remains sensitive. So long as the underlying bilateral issue remains unsettled, the risk to the global semiconductor supply chain continues.

US-China frictions

Semiconductors occupy a prominent position in the ongoing tensions between the US and China that have escalated significantly since 2018. While semiconductors have been largely excluded from tariffs that both countries enacted on a range of imports from the other side, in 2019 and 2020 the US government imposed a series of export controls that restrict access to semiconductors containing US technology for Huawei and other Chinese entities that it regards as acting contrary to US national security or foreign policy interests.

As of March 2021, some of these export controls encompass the entire semiconductor supply chain, including EDA and manufacturing equipment that incorporates technology developed in the US. Given that US companies are currently the only viable suppliers of EDA and critical equipment such as doping or metrology (see Exhibit 17), these controls for now effectively block the impacted Chinese entities from sourcing semiconductors, even from non-US suppliers. These rules have encouraged China to develop and seek alternatives, and although it may take some time to do so, the trend towards reduction of dependence on US semiconductor suppliers and indigenization of the supply chain is beginning to take shape.

As described in the prior section, China accounts for approximately 24% of the global semiconductor consumption (“criteria C” in Exhibit 3), which makes it the second largest market in the world almost at par with the US. Its position as the world’s largest manufacturing hub for electronic devices – for both Chinese and foreign companies – also makes China the top destination for exports of finished chips. In addition, China is investing aggressively in semiconductor manufacturing: it accounted for 15% of the world’s total capacity in 2020 and is forecasted to build 40% of the incremental capacity that will be added globally in the next decade.

Geopolitical tensions could lead to the loss of global scale required to fund massive investments in innovation

Continuation of these bilateral tensions could have profound negative consequences for the semiconductor industry. Both US semiconductor companies, and also foreign vendors that rely on technology developed in the US, may be blocked from selling to at least some significant Chinese customers, if not to any Chinese company at all. As discussed in our March 2020 report on this topic, this could lead to a significant reduction in revenue for leading US semiconductor companies across the supply chain as well as global non-US companies with a significant R&D footprint in the US, compromising their ability to sustain their current investment levels in R&D and therefore slowing the pace of innovation across the industry.

Perpetuation of the conflict may also trigger retaliation from China in areas that could directly or indirectly impact the semiconductor supply chain, such as rare earth materials and ten other critical inputs such as germanium, lithium or tungsten. Rare earths are a set of 17 metallic elements with electronic and magnetic properties needed in electronic products. Although these materials account for only a small portion of overall production costs, they are the building blocks of key components in cars, computers, and many other high-value products—and are an often-overlooked vulnerability in global supply chains. Our analysis indicates that China leads in the extraction of 9 of the 17 critical raw-material inputs and in the refining of 14 of them. As rare earths are traded in commodity markets, restrictions on exports from China would be felt by the entire supply chain and could disrupt the global production of electronic devices and therefore depress demand for semiconductors.

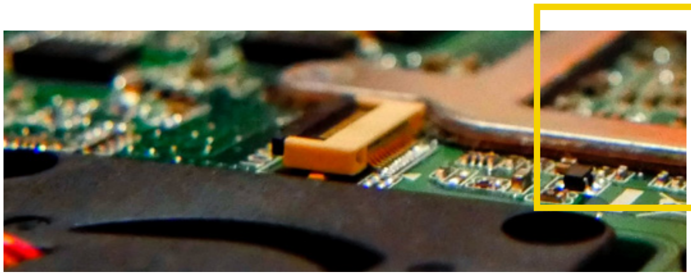
Finally, the US-China frictions are also fueling a desire to develop self-sufficiency in semiconductors. For China, this is mainly an amplification and acceleration of its longstanding efforts to develop a strong domestic semiconductor sector that gained further urgency with the “Made in China 2025” plan introduced in 2015.

In the case of the US, the escalating strategic competition with China has recently exposed the risks associated with the high concentration of semiconductor manufacturing capacity in East Asia (and Taiwan in particular), sparking some public debates about the desirability of self-sufficiency in semiconductor manufacturing, too.

In other areas of the world such as Europe, Japan and South Korea, the central position of semiconductors in the US-China conflict together with the impact of the recent widespread semiconductor shortage on the automotive industry, has brought attention to the critical importance of semiconductors for the economy. Furthermore, their own companies with global leading positions in some segments of the semiconductor industry have found themselves restricted from selling to Chinese entities by US export controls due to their reliance on US-developed technology further upstream or downstream in the value chain.

Addressing these risks: complete self sufficiency is not the answer

Semiconductors are of strategic importance for both economic growth and national security. The semiconductor supply chain has become a critical area whose operational resilience and continuity must be enhanced, as well as a heated field of geopolitical competition for the 21st century.



In view of the two risks described above, governments across the world are looking to act. The concepts of semiconductor “self-sufficiency”, or technology “independence” or “sovereignty”, are being discussed as potential desirable national policy goals – often with a focus on semiconductor manufacturing. It is helpful to understand what level of investment would be needed if most countries or regions were to re-shore or nearshore production capacity to reduce exposure to these risks and protect their national interests. We look at two scenarios – one where each region pursues complete semiconductor self-sufficiency compared to more nimble, targeted investments aimed at filling strategic high-risk gaps to improve resilience in the overall global supply chain.

For illustration purposes, Exhibit 20 presents a hypothetical extreme scenario, where each major region in the world looks to build up semiconductor “self-sufficiency” in a strict sense, across all layers of the supply chain. This would mean having domestic firms in EDA and IP cores, chip design, raw materials, manufacturing equipment, wafer fabrication, and assembly, packaging and testing, with enough capacity to meet 100% of the domestic semiconductor consumption across all applications.

Aside from any considerations of execution feasibility, we estimate that at a global level such extreme scenario of regional autarchy would require a staggering $900 to 1,225 billion in upfront investment to cover each region’s 2019 consumption levels – any future growth in domestic consumption would require further investments in additional capacity by each region. This amount is equivalent to about 6 times the combined R&D investment and capital expenditure of the total semiconductor value chain in 2019. In addition, even if we were to assume that semiconductor companies across the supply chain could maintain their current cost structure despite the loss of global scale, we estimate that the industry would incur $45 to $125 billion in incremental recurrent annual operational costs (Exhibit 20)[15].

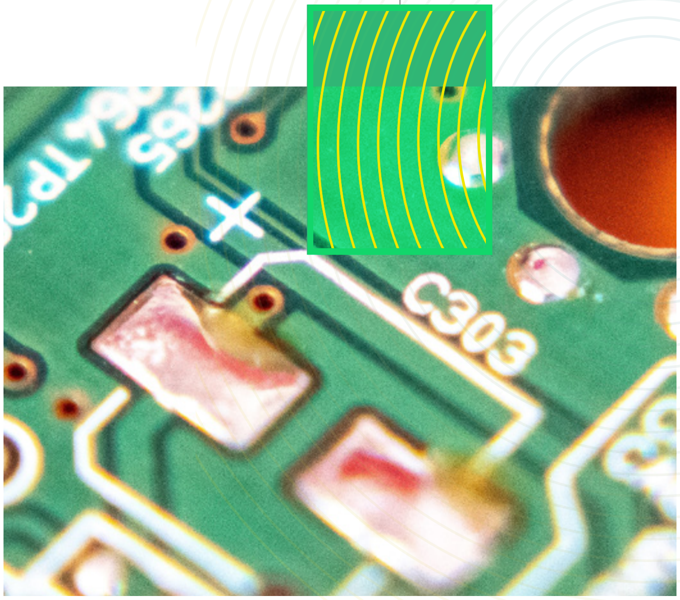
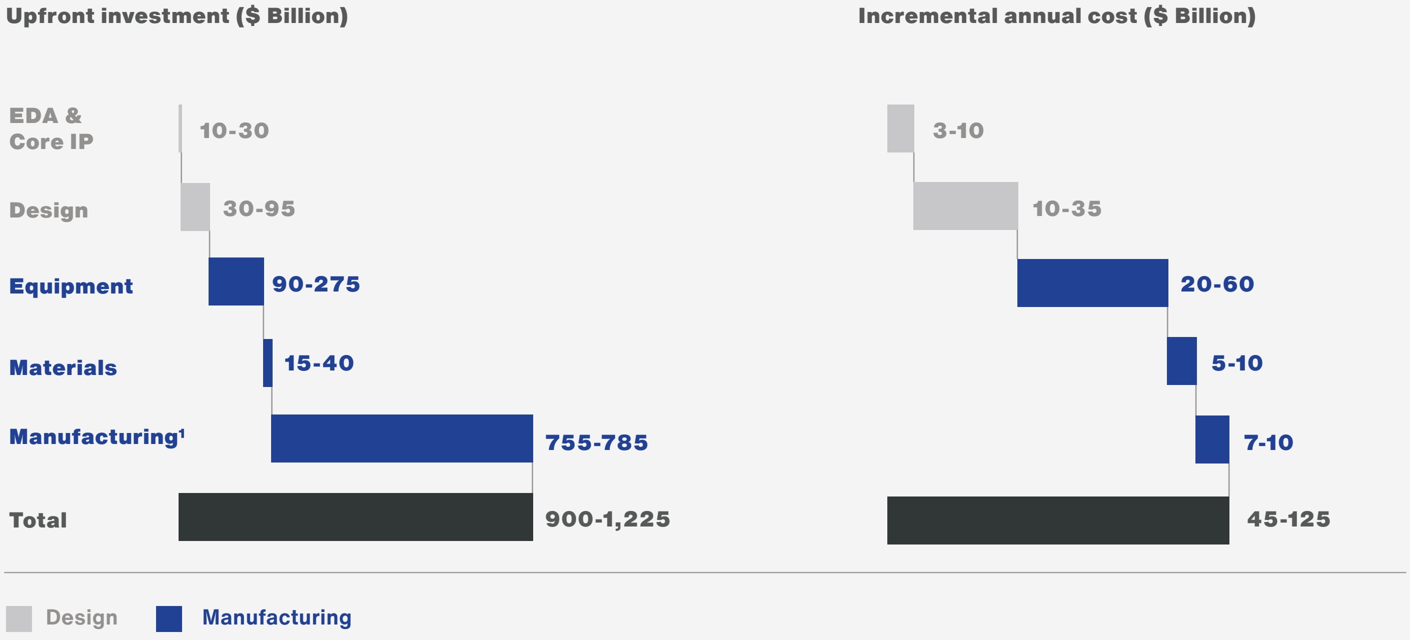


EXHIBIT 20

The staggering cost of hypothetical semiconductor self-sufficiency

Incremental cost to cover 2019 demand with fully "self-sufficient" localized semiconductor supply chains



1. Including both wafer fabrication and assembly, packaging and testing
Note: Range defined primarily by number of local companies assumed to be required to meet the local needs in each activity of the value chain: from just 1 player to supply the entire local market to 3 players typically found in the current global market structure
Sources: BCG analysis

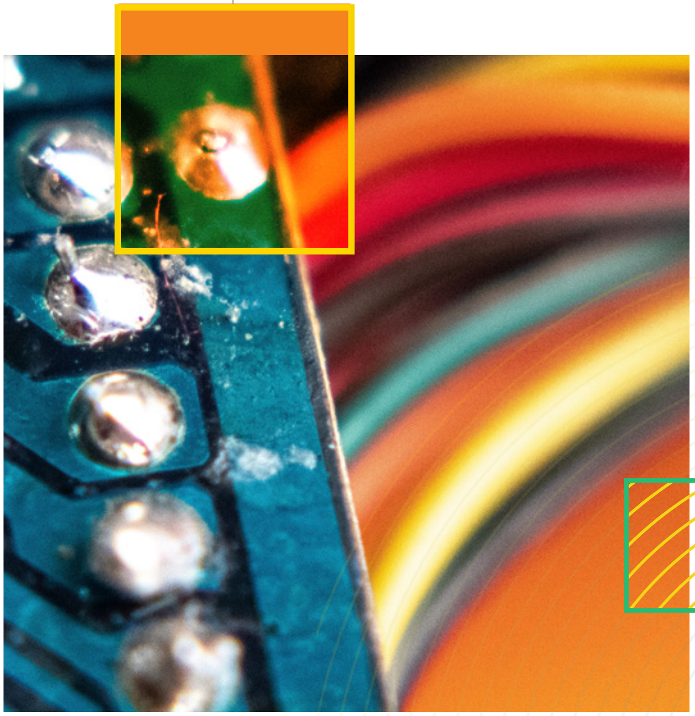
---

[15] Range defined primarily by number of local companies assumed to be required to meet the local needs in each activity of the value chain: from just 1 player to supply the entire local market to 3 players typically found in the current global market structure

Manufacturing capacity accounts for a majority of the estimated upfront investment. In total, the new capacity that would have to be built in this extreme scenario is equivalent to 40% of the total existing global capacity – which suggests that the industry would face a situation of massive global overcapacity if such scenario were to be implemented. We estimate that over $800 billion in capital expenditure would be required to build the local raw material production, wafer fabrication and assembly, packaging and testing capacity needed for each region to meet its own domestic semiconductor consumption. Obviously, in addition to securing the funding for these massive investments, countries or regions would also have to actually make it happen: fabs are very complex specialized facilities that typically take 2-4 years to build and put into commercial production, and require 3,000 to 6,000 staff to operate, mostly skilled technicians that need to be recruited and trained.

On top of this upfront investment in new capacity, each country or region would incur the costs to operate the fabs. Even assuming that over the long term the industry is able to eventually reestablish the supply-demand balance and avoid a continued state of global overcapacity, the total global manufacturing operating costs would significantly exceed the levels of the current global supply chain structure with geographic specialization based on comparative advantages. As discussed in the prior section on the benefits of a globalized value chain, in this hypothetical scenario of manufacturing self-sufficiency, there will be capacity located in countries with factor costs (such as land, labor and electricity) significantly higher than those available in other geographic areas. We estimate that the overall excess manufacturing cost could amount to $7-10 billion per year globally – not including the effect of potential differences in government incentives and taxation, or the depreciation of the upfront capital investments.

However, self-sufficiency in manufacturing would not totally eliminate the exposure to the two risk factors described above, which also affect the other layers upstream and downstream in the semiconductor supply chain. If self-sufficiency were to be achieved in semiconductor design, each region would need to replicate its own competitive domestic supplier for each of the over 30 types of semiconductors described earlier in this report. Furthermore, each region would need to replicate domestically its own EDA and core IP, as well as its own manufacturing equipment. We estimate that this will add $130-400 billion of upfront investment to cover between 5 and 15 years of start-up period R&D to develop the local technology prior to commercialization. These figures do not take into account any potential failed investments, such as companies that start R&D but do not succeed in developing a commercially viable product. In addition, the new local firms involved in these activities would face $33-105 billion in incremental annual operating costs – mainly in recurrent R&D.



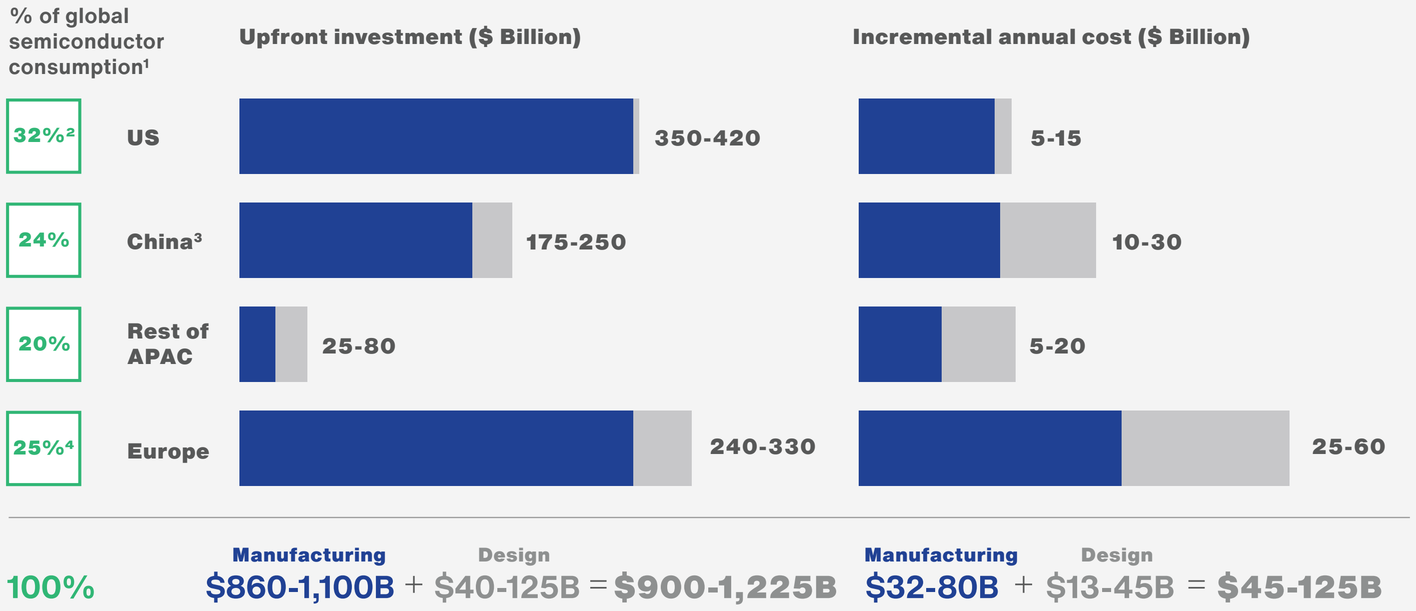
Fully “self-sufficient” localized supply chains would create substantial incremental costs and lead to a 35-65% increase in semiconductor prices

Exhibit 21 provides the breakdown of the estimated cost of hypothetical semiconductor self-sufficiency by geographic area. The US, despite its current leadership position in several layers of the supply chain, would still need to make an upfront investment of $350-420 billion, primarily in manufacturing capacity. Even considering its lower factor costs, China would require $175-250 billion of upfront investment and $10-30 billion of additional incremental annual operating costs in this hypothetical state.

EXHIBIT 21

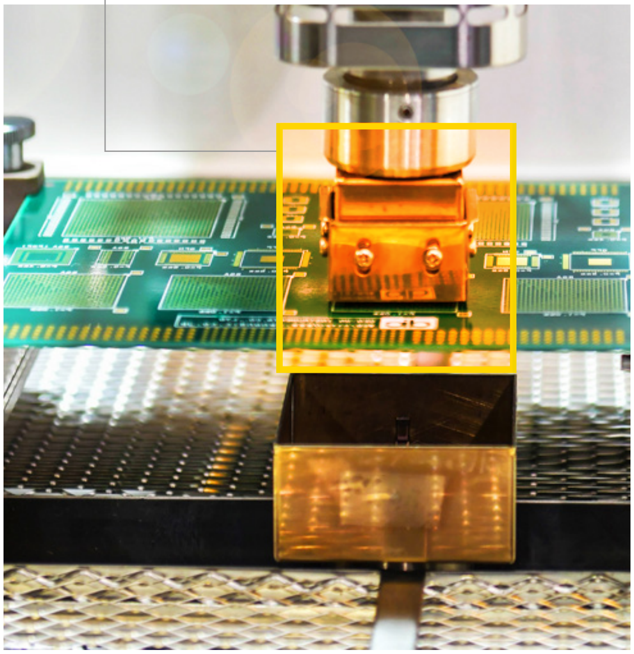
All regions benefit from the efficiencies of the global semiconductor supply chain

Incremental cost to cover 2019 demand with fully "self-sufficient" localized semiconductor supply chains



1. Calculated as the estimated semiconductor content in electronic devices sold to end users in each geography in 2019 2. Includes Canada, Latin America 3. Mainland China 4. Includes Middle East and Africa
Note: Range defined primarily by number of local companies assumed to be required to meet the local needs in each activity of the value chain: from just 1 player to supply the entire local market to 3 players typically found in the current global market structure
Sources: BCG analysis

This total estimated $900-1,225 billion in upfront investment and $45-125 billion in incremental annual operating cost – not including depreciation of the new upfront investments - to build a series of fully domestic supply across each major region would all but wipe out the profits of the industry, which amounted to $126 billion across the entire value chain in 2019. Therefore at least a portion of the incremental costs would necessarily have to be passed on to device makers in the form of higher prices for the semiconductors they purchase. If fully charged to customers, it would amount to an average increase of 35-65% in the price of semiconductors. This may result in higher prices of the electronic devices for end users. Furthermore, it is also likely that siloed domestic industries shielded from foreign competition and deprived of global scale would lose in efficiency and ability to innovate. Ultimately, it would reverse the decades-long trend of making increasingly powerful and more affordable electronic devices accessible for consumers around the world.



In conclusion, complete autarchy or full semiconductor self-sufficiency appears to be more a theoretical concept rather than an attainable policy goal. However, countries can reaffirm their position within the global supply chain by ensuring they contribute to technology development with research and intellectual property. This ensures mutual dependency and provides each nation with a position of strength.

An alternative, market driven approach focused on critical strategic risks

Countries with serious economic and national security concerns may take a more focused approach, building some advanced semiconductor manufacturing capacity domestically in order to address their most sensitive needs in critical application areas.

In our **previous report** titled "Government Incentives and US Competitiveness in Semiconductor Manufacturing," we found that the US share of semiconductor manufacturing capacity has dropped from 37% in 1990 to 12% in 2020. Moreover, only 6% of the new global capacity in development will be located in the US if current trends continue. In contrast, it is projected that during the next decade China will add about 40% of the new capacity and become the largest semiconductor manufacturing location in the world. As discussed earlier, the key factor behind this trend is economics: the total ten-year cost of ownership of of a new fab located in the US is approximately 25-50% higher than in Asia, and 40-70% of that difference is attributable directly to government incentives (see Exhibit 16).

According to our prior analysis a $20 billion to $50 billion federal government program of additional grants and tax incentives for new state-of-the-art fabs built in the next decade would be effective in beginning to reverse the declining trend in US semiconductor manufacturing of the last 30 years. For example, we estimate that a $50 billion incentive program would enable the construction of 19 fabs in the US over the next ten years, doubling the number expected if no action is taken. In contrast, a goal of complete self-sufficiency to cover the total US semiconductor consumption by 2030 with onshore capacity would require more than $400 billion in government incentives.

Exhibit 22 describes how this market-driven, targeted $20-50 billion government incentive package would, for example, enable the US to maintain "minimum viable capacity"[16] for advanced logic chips, essential for national security and supply chain resilience. While all types of semiconductors from memory, logic, to analog are indispensable, each performing a different function to make electronic devices work, advanced logic chips – including CPUs, GPUs, FPGAs, AI accelerators and ASICs – have attracted particular focus. Advanced logic chips rely on manufacturing on the leading nodes to maximize performance. As shown in Exhibit 17 earlier, currently all the world's capacity below 10 nanometers is located in South Korea (8%) and Taiwan (92%). Given the importance of advanced logic chips for technology leadership in high-performance computing and AI, the US has recently identified it as a vulnerability in the microelectronics supply chain that poses a potential national security risk[17].

Advanced logic chips account for about 34% of US total semiconductor consumption. A significant portion of that figure actually comes from consumer-

---

[16] This concept was introduced by the White Paper #4 from the US Cyberspace Solarium Commission published in October 2020.

[17] See for example the 2021 final report of the National Security Commission on Artificial Intelligence (NSCAI)

driven applications, such as smartphones, PCs, consumer electronics and automotive. However, about a quarter of the US consumption of advanced logic chips is associated with critical infrastructure applications, including aerospace and defense systems, core telecommunications networks, supercomputers and data centers for essential sectors such as government, energy, transportation, healthcare and financial services. A hypothetical disruption in the supply of these chips could have a severe impact on the economy and national security, so maintaining some minimum viable manufacturing capacity located onshore could significantly enhance the resilience of the US electronics supply chain. Covering just 9% of the total US semiconductor consumption, such targeted intervention is far removed from a large-scale industrial policy aimed at building a self-sufficient local semiconductor supply chain.

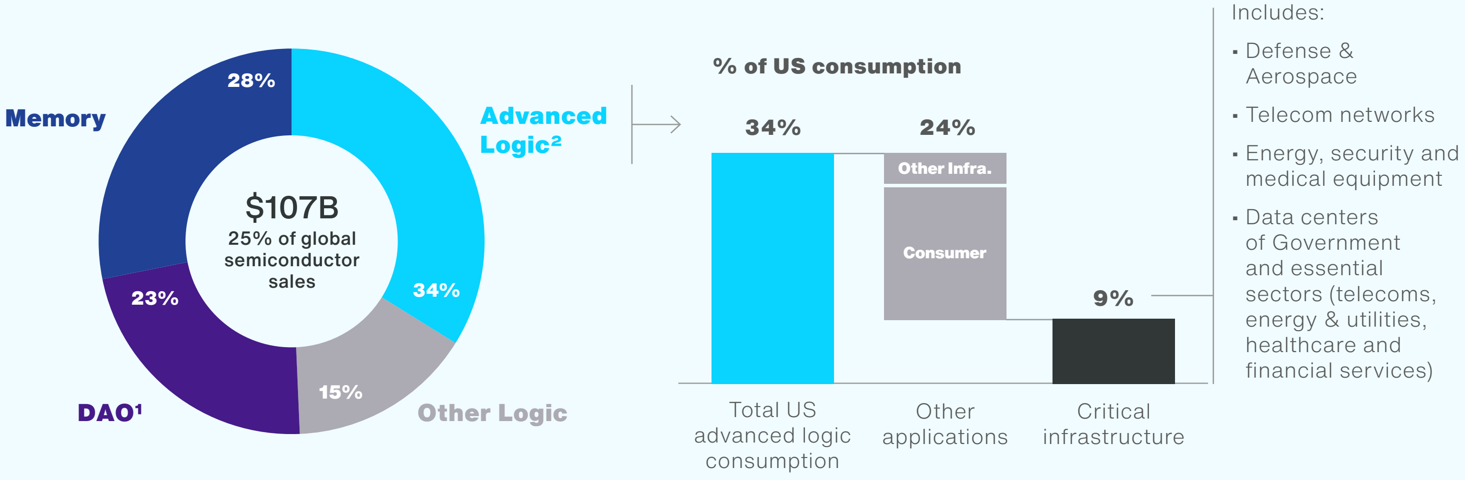
We estimate that covering the expected domestic consumption of advanced logic chips for critical infrastructure applications by 2030 would require building just 2-3 new state-of-the-art fabs in the US – assuming new fabs with capacity between 20,000 and 35,000 wafers per month. This capacity addition is well aligned with market needs. In fact, this new capacity is less than 5% of the new advanced logic capacity that needs to be added globally to keep up with the expected demand growth in the next ten years. The presence in the US of existing advanced manufacturing infrastructure and most of the world's leaders in design of advanced logic chips, which account for a large share of foundry revenues, is also a relevant factor that contributes to the viability of such investment.

The new $20-50 billion government incentive program mentioned above would be critical to make the economics of such new advanced logic fabs in the US competitive with alternative available locations in Asia. We estimate that these fabs would require a $40-45 billion of private sector investment to build and operate over a 10-year period, together with $15-20 billion of government incentives—of which $9-10 billion would need to come from the new federal incentive program. The remaining new federal incentives would be applied to spur investment in new US capacity in other important areas besides advanced logic, including memories, analog and advanced packaging.

EXHIBIT 22

Example: How the US could establish a minimum viable capacity to address its critical strategic risk in advanced logic manufacturing

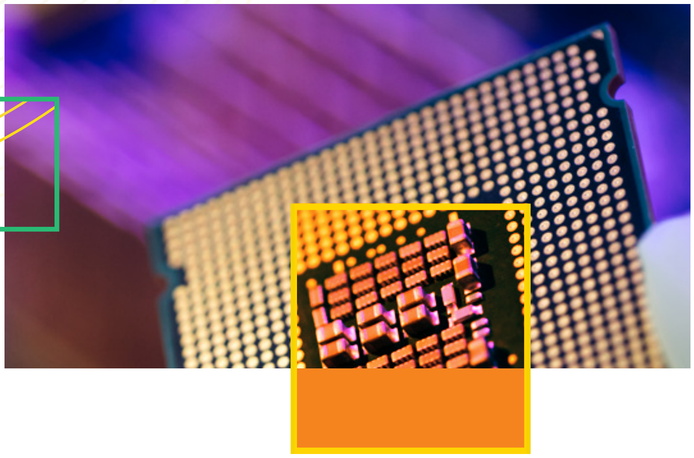
US ANALYSIS: Breakdown of total US semiconductor consumption, 2019



% US Capacity / Consumption	49%	0%[3]	0%[3]
10-year investment in new fabs for onshore coverage in 2030 ($B) (Private sector + Government incentives)[4]	980+420	140+60	45+18 2-3 new fabs of 20-35 kwpm

1. Discrete, Analog and other (optoelectronics and sensors) 2. Advanced logic includes CPUs, GPUs, Application Processors, FPGAs, mobile basebands, some ASICs 3. Considering only leading node capacity (< 10nm) 4. Total Cost of Ownership – includes capex and 10 years of opex, before government incentives

Sources: BCG analysis



Altogether, these considerations – critical strategic area for technology leadership and national security, geographic proximity of foundries to their large US-based advanced logic customers, moderate size of the investment – make this an attainable policy goal. In fact, according to media reports the three companies in the world that are currently capable of manufacturing at 10 nanometers or below (TSMC, Samsung, and Intel) are showing interest in building some of their planned additional advanced logic capacity in the US, provided that there is a package of government incentives that make the economics work.

In setting policies aimed at advancing the domestic semiconductor sector to address supply chain vulnerabilities and national security concerns, governments should ensure that they take a market-driven approach. Government policies and incentives must also be in compliance with the norms of international trade. The OECD report[18] on market distortions in the semiconductor supply chain from 2014 to 2018 was already drawing attention to some government support practices – particularly from China – that could be considered anti-competitive, or lead to inferior economic outcomes. Government support to the build-up of the domestic semiconductor manufacturing sector should also take into account the expected global needs for new capacity to meet the expected growth in worldwide demand to avoid creating situations of massive overcapacity.

The talent constraint

Access to high-skilled talent is also critical for an R&D-intensive industry like semiconductors. Unlike the other two risks described above, a shortage of talent may not pose an immediate threat of large-scale disruption for the industry day-to-day operations. Nonetheless, it can significantly impair its ability to continue its rapid, relentless pace of innovation in the upcoming years, and the diversification of the global geographic footprint in some activities of the supply chain.

Indeed, talent has become a major concern for the industry. A 2017 survey of semiconductor executives across the supply chain showed that about 80% of companies were facing significant shortages of candidates for technical roles[19]. In another 2018 survey 64% of respondents named talent as one of the top 3 risks threatening their ability to grow, the highest identified risk factor[20]. Salary statistics also point to talent supply constraints: wages in the US semiconductor industry have been growing an average of 4.4% since 2001, significantly faster than the growth in wages for the economy as a whole[21].

Considering that the percentage of R&D investment over revenues tends to be quite stable over the medium term for many semiconductor companies, growth in industry revenues may be a good proxy to set a baseline for the increasing demand for talent. Global semiconductor industry sales are expected to increase at a 4-5% average annual rate in this decade. On top of this growth, the industry is also facing the challenge of an aging workforce, with a significant number of current employees in technical positions likely to retire in the next 10-15 years. Furthermore, the industry also needs to attract talent with different skill sets, particularly in software development and artificial intelligence.

The industry faces a risk of talent shortage that could constrain the pace of innovation in upcoming years

The historical growth in the total global talent pool of science and engineering graduates looks insufficient to meet the industry demand for talent (Exhibit 23).

---

[18] OECD, "Measuring distortions in international markets: The semiconductor value chain", OECD Trade Policy Papers, No. 234, OECD Publishing, Paris, 2019.
[19] Deloitte-SEMI Workforce Development Survey, 2017.
[20] KPMG-GSA, Global Semiconductor Industry Outlook 2019.
[21] Semiconductor Industry Association, SIA Workforce Roundtable Summary Report, March 2018

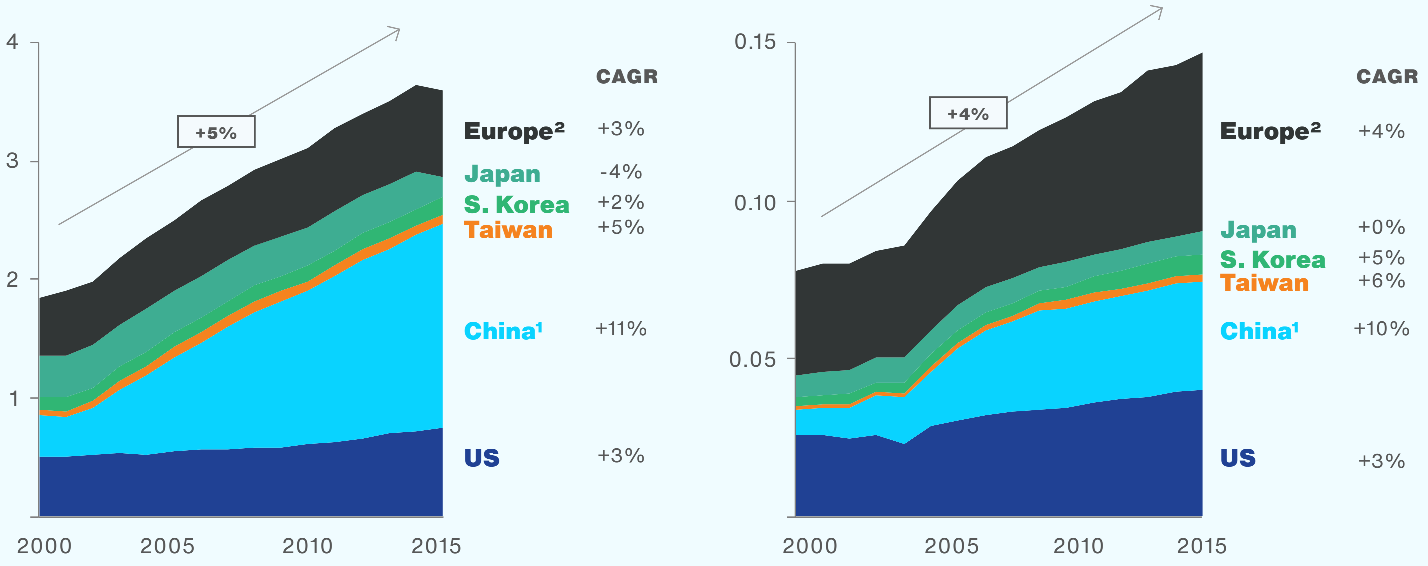
Our analysis of the global data compiled by the US National Center for Science and Engineering Statistics (NCSES) shows that the aggregate number of students with a first university degree in science and engineering in the leading regions in the semiconductor supply chain - US, mainland China, Taiwan, Japan, South Korea and Europe - grew at 4.5% annually between 2000 and 2015 (most recent year with complete data). The number of doctorates in science and engineering showed a very similar growth rate. This growth was also quite different across regions: while China's talent pool grew above 10% per year, in the US the growth rate was below 3%.

There is also fierce competition for this global talent pool. In particular, the explosion in the number of software and consumer technology companies, including global giants with well-known brand names, adds to the industry challenges in attracting and retaining the high-skilled technical talent it needs to maintain the current trajectory in innovation and growth.

EXHIBIT 23

The historical growth rate of the total global talent pool is likely insufficient to meet the industry demand for talent

Annual graduates in Science and Engineering (millions)



First university degrees

Doctoral degrees

1. Mainland China
2. Top 5 countries
Surces: BCG analysis with data from US National Center for Science and Engineering Statistics (NCSES)

Strengthening the Global Supply Chain for the Next Decade of Innovation

The global nature of the semiconductor supply chain provides specialized firms in each layer with access to customers across all markets, which brings the scale needed to fund the large investments in R&D or manufacturing capacity to design and produce semiconductors. It has also enabled geographic specialization based on the comparative advantages of different countries or regions: the US is the leader in the most R&D-intensive parts of the value chain such as chip design and manufacturing equipment, while Asian countries with access to higher government investment incentives and lower factor costs lead in manufacturing. Within this global supply chain, firms collaborate and trade semiconductor-related goods and services across borders, and countries are mutually interdependent.

Looking ahead, over the next ten years the industry will need to invest about $3 trillion in R&D and capital expenditure across the value chain to meet the fast-growing demand for semiconductors from all sectors of the global economy. Semiconductor companies will need to keep investing over $90 billion in R&D annually, equivalent to about 20% of the global semiconductor sales, to develop increasingly sophisticated chips to power transformative applications such as AI, IoT or autonomous vehicles.

On the manufacturing side, as the global shortage of semiconductors of late 2020 is highlighting, a substantial amount of new capacity needs to be built globally in the coming years. In fact, the industry will have to almost double its capacity by 2030 to keep up with the expected 4% to 5% average annual growth in semiconductor demand.

Simultaneously, the industry must innovate in materials, architectures and manufacturing technology if the rate of improvement in performance and cost of the past few decades is to be maintained in years to come. This requires investing heavily in pre-competitive basic research.

A strong global supply chain that continues bringing world-class firms together to collaborate on innovation in materials, design and manufacturing across borders is vital to make this possible. At the same time, the identified risks related to the high degree of geographic concentration of some critical parts of the supply chain such as wafer fabrication, as well as the potential disruptions from geopolitical frictions must be addressed to make the semiconductor global chain more resilient.

The solution to these challenges is not to pursue blanket self-sufficiency through large-scale national industrial policies that come with a staggering cost and questionable execution feasibility. Instead, we believe that well-modulated policy interventions in these areas would simultaneously preserve the benefits of scale and specialization in today's global supply chain structure, strengthen its resilience, and also address the national security concerns associated with the strategic nature of semiconductors.

> In addition to policies that foster trade, basic research and expand the talent pipeline, governments need to enact targeted incentives to support investments that diversify the global manufacturing footprint and the sources of supply for key materials

To that end, policies that expand market access and promote open trade while also balancing the needs of national security are fundamental to allow the complex semiconductor industry ecosystem to continue to thrive in the next decade. First, governments must guarantee a level global playing field with market access in fair terms for domestic and foreign firms alike, as well as strong protection for IP rights. In fact, such policies could further encourage foreign investment in R&D activity, favoring inflows of know-how and talent that ultimately help upgrade the capabilities of the domestic industry and stimulate healthy competition in innovation and quality.

Second, governments should seek to develop policies to preserve and expand the global trade in semiconductor-related goods and services that underpins the geographic specialization based on comparative advantage in the semiconductor supply chain. The explosion in trade enabled by the World Trade Organization's Information Technology Agreement (ITA) in the last two decades enabled the emergence of the global supply chain and fuelled the industry's growth, so its expansion is a clear positive development for the industry.

While safeguarding national security interests is of course critical, policy mechanisms require careful consideration if they are to avoid permanently harming the specialization model that has enabled the semiconductor industry's success. Our prior March 2020 report focused on the potential implications of the US-China frictions for the semiconductor industry showed that broad unilateral restrictions on access to markets, technologies and resources may backfire and risk endangering the the long-standing global leadership of the US in semiconductors, with a detrimental impact on the industry's innovation, too.

Instead, governments with significant national security concerns related to control over semiconductor technology should establish a stable framework for restrictions on semiconductor trade that defines with clarity:

a) the policy goals pursued (control over sensitive military technology, reciprocal fair trade, strategic competition in certain technology areas, retaliatory sanctions related to other conflict areas);

b) which types of entities and specific technologies are restricted;

c) the expected second-order impacts on industry players, both domestic and from third countries, and what measures to put in place to mitigate them.

Finally, policy makers must significantly step up the efforts to address the looming shortage of high-skilled talent that threatens to constrain the semiconductor industry's ability to keep the current pace of innovation and growth. Governments should invest further in expanding science and engineering education across an array of disciplines including electrical and mechanical engineering, computer science and software engineering, and physics, materials science and chemical engineering, as well as industrial engineering. Investment is also needed to foster the creation of more semiconductor-related graduate programs, together with partnerships between government agencies, academic institutions and the industry to set up programs that can accelerate training in specific technical areas. In parallel, immigration pathways that attract world-class technical talent to leading global clusters of semiconductor innovation should be maintained and reinforced.

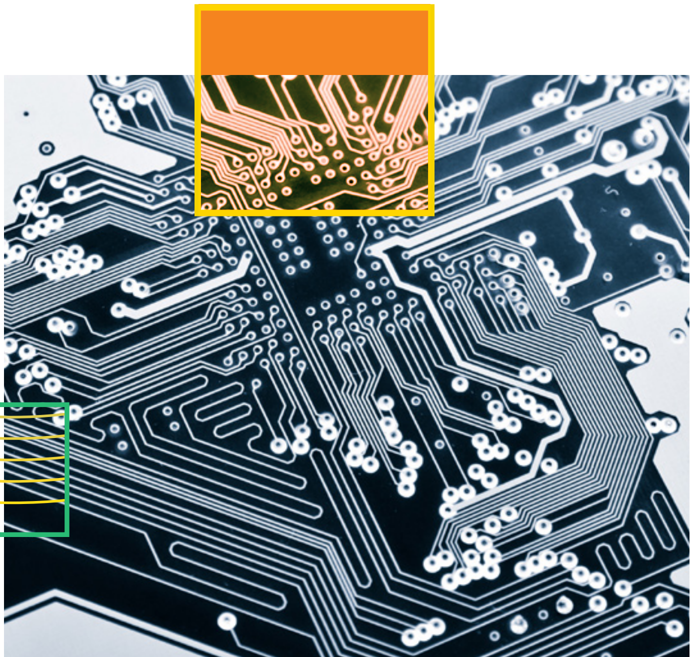
As shown in this report, these policy directions should be complemented with targeted, market-driven government incentives programs in three areas:

- Pre-competitive research, which is often channelled through global institutes that foster collaboration between scientists from different organizations and countries

- Building part of the additional manufacturing capacity that the industry needs to meet demand in the next decade in the US and Europe, in order to create a more diversified global footprint that is less vulnerable to disruption by natural disasters, infrastructure failures, cyberattacks or geopolitical frictions.

- Developing alternatives – domestic or from third countries—in critical areas threatened by export controls or potential disruptions to global trade from other countries

In our view, carefully designed policy initiatives in these areas are essential to enable the industry to sustain its high levels of investment in both R&D and capital expenditure to meet the fast-growing demand for higher quantities of more advanced semiconductors that the global economy needs to make the promise of transformative technologies such as AI, 5G, IoT, and autonomous electric vehicles real in the upcoming years.



BCG × SIA SEMICONDUCTOR INDUSTRY ASSOCIATION